

1984

STUDIES ON THE STRUCTURE AND ORGANIZATION OF THE SOYBEAN CHLOROPLAST GENOME, CLONING AND LOCALIZATION OF GENES.

GURPARKASH. SINGH

University of Windsor

Follow this and additional works at: <http://scholar.uwindsor.ca/etd>

Recommended Citation

SINGH, GURPARKASH., "STUDIES ON THE STRUCTURE AND ORGANIZATION OF THE SOYBEAN CHLOROPLAST GENOME, CLONING AND LOCALIZATION OF GENES." (1984). *Electronic Theses and Dissertations*. Paper 3645.

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.

CANADIAN THESES ON MICROFICHE

I.S.B.N.

THESES CANADIENNES SUR MICROFICHE



National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Ottawa, Canada
K1A 0N4

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

Studies on the Structure and Organization of the Soybean
Chloroplast
Genome ,Cloning and Localization of Genes

by

GURPARKASH SINGH

A Dissertation
presented to the University of Windsor
in partial fulfillment of the
requirements for the degree of
Ph. D.
in
Department of Biology

799092

Windsor , Ontario, 1983

(c) GURPARKASH SINGH, 1983

ABBREVIATIONS

AMP	Adenosine 5' monophosphate
ATP	Adenosine 5' triphosphate
dATP	Deoxyadenosine 5' triphosphate
dCTP	Deoxycytidine 5' triphosphate
dGTP	Deoxyguanosine 5' triphosphate
dTTP	Deoxythymidine 5' triphosphate
BME	2- Mercaptoethanol
BSA	Bovine serum albumin
bp	base pairs
Kbp	Kilobase pairs
Ci	Curie
cpm	Counts per minute
cpDNA	Chloroplast DNA
DCCD	NN' Dicyclohexylcarbodiimide
DEAE-cellulose	Diethylaminoethyl-cellulose
DNA	Deoxyribonucleic acid
DNAase	Deoxyribonuclease
EDTA	Ethylenediamine tetra acetic acid
IPTG	Isopropyl thio - β -galactoside
μ g	microgram
μ l	microliter
μ m	micrometer
ng	nanogram
nm	nanometer
RNA	Ribonucleic acid
RNase	Ribonuclease
RPC	Reversed phase chromatography
rpm	revolutions per minute
Krpm	1000 revolutions per minute
SDS	Sodium dodecyl sulphate
SSC	Standard Saline citrate (150 mM NaCl, 15 mM Sodium citrate)
TEMED	Tetramethylethylenediamine
V/V	Volume by volume
W/W	Weight by weight
Xgal	5 bromo-4-chloro-3-indolyl- β -D galactoside

To My Parents

ABSTRACT

The chloroplast DNA from soybean (Glycine max L.) was isolated and characterized. Cleavage sites of restriction endonucleases KpnI, Pvu II, Sac I were arranged with respect to endonuclease Xho I sites by filter hybridization using Xho I fragments as labeled probes. A total of 66 sites were aligned on the soybean chloroplast genome. Restriction analysis indicates a circular DNA molecule of 151 Kb which is organized into two unequal single copy regions of 84 Kb (large) and 20 Kb (small). The large and small single copy regions are interrupted by a single 23.5 Kb repeat region on either side. The repeated region is organized in an inverted orientation.

Position of rRNA genes was determined by hybridizing radio-labeled 16S and 23S rRNAs. Results indicate that rRNA encoding region is present once per segment of the inverted repeat, hence two copies are present per chloroplast genome. The two sets of rRNA genes are separated by approximately 20 Kb in one direction and 116 Kb in another.

In order to undertake detailed fine mapping, structural and sequence organization studies, it was necessary to raise a clone bank of soybean chloroplast DNA. The entire genome was cloned in 4.36 Kb pBR322 and 2.5 Kb pDPL13. A set of seven Pvu II fragments and a set of eight Sac I fragments were inserted in pBR322 and pDPL13 vectors respectively. Both insertions were outside the drug markers.

Fine mapping of the cloned fragments containing the rRNA cistron (rDNA) suggests the presence of a single copy of 16S and 23S rRNA genes per rDNA unit. The 16S and 23S sequences in each unit are separated by 2.3-3.0 Kb spacer. The direction of transcription is from 16S to 23S.

The tRNA genes for leucine, (trnL1, trnL2, trnL3), serine (trnS), phenylalanine (trnF) were localized by Southern hybridization.

Plastid protein encoding genes were mapped by hybridizing internal region of gene probes from other plants to soybean restriction digests. The genes for ribulose-1,5-bisphosphate (rbcl), alpha, beta and epsilon subunit of CF₁ (atpA, atpB and atpE respectively), Subunit III (DCCD binding polypeptide) of CF₀ of ATP synthase (atpH), the cytochrome f (CytF), and the "32Kd" polypeptide (psbA) are located in the large single copy region.

ACKNOWLEDGEMENTS

I take this opportunity to sincerely thank Dr. D.T.N. Pillay, my adviser, for guidance, support, and research facilities during the course of this work. I also thank him for introducing me to recombinant DNA research and giving me the opportunity to accomplish this task. Acknowledgements are also due to his wife, Sarah, for her kindness and generosity.

My appreciations are extended to other members of my research committee; Drs. D.A. Cotter, N.F. Taylor, and D. des S. Thomas for encouragement and suggestions in preparation of this manuscript. Thanks are also forwarded to Prof. J.H. Weil for consenting to serve as external examiner, and for reviewing the work on such short notice.

Many thanks to Dr. J.D. Palmer for patiently teaching me recombinant DNA techniques, and for supplying various gene probes. Thanks are also extended to Drs. E. Crouse and A. Steinmetz for stimulating discussions.

I extend my gratitude to Dr. N.A. Straus for direction, technical assistance, and philosophical discussions during my stay in his laboratory. Also, sincere thanks go to Dr. Straus and his wife, Leslie, for accepting me and cheerfully providing many comforts during my stay in their home.

I thank the following scientists for their generous gifts; Dr. D. Bourque for ribosomal RNA probes; Drs. H.J. Bohnert, J.C. Gray and R.G. Herrmann for chloroplast gene probes; Dr. D. Pulleyblank for the plasmid, pDPL13; Dr. R. Gege for yeast nucleotidyl transferase and Dr. P. Guillemaut for tRNAs.

Next, I wish to thank my friends, especially Magdalena Skunca for reassurance during the tough moments; Guy de Lanversin for the constructive criticism and friendship; K. Suryanarayana for timely advices; Diwakar Gupta, Mohan Krishnan, Chandershekhar and Raj Patil for assistance with Syspub.

Lastly, I wish to thank Sandy Dunlop for assistance in developing autoradiograms; Margie Prytulak for cheerfully working on computer, even at odd hours to meet my deadline, and Sherry Crooks for drawings.

I acknowledge the receipt of a University of Windsor Graduate Scholarship.

This research was funded by NSERC grant to Dr. D.T.N. Pillay and in part by another NSERC grant to Dr. N.A. Straus.

CONTENTS

ABSTRACT	iii
ACKNOWLEDGEMENTS	iv

<u>Chapter</u>		<u>page</u>
I	INTRODUCTION	1
	Presence of extra nuclear genetic system	1
	Physicochemical properties of chloroplast genome	2
	Physical structure and organization	3
	Cloning	7
	Protein encoding genes	8
	Purpose of the project	9
II	PHYSICAL MAPPING OF SOYBEAN CHLOROPLAST	11
	INTRODUCTION	11
	MATERIAL AND METHODS	13
	Enzymes and supplies	13
	Chloroplast isolation	13
	Isolation of DNA	14
	Cleavage of chloroplast DNA by restriction endonuclease	15
	Gel electrophoresis	16
	Staining and photography	16
	DNA images	16
	Nick translation	17
	Hybridization	17
	Washing	18
	Autoradiography	18
	RESULTS	19
	Selection of endonucleases	19
	Molecular weight determination	19
	Mapping Strategy of Cleavage Sites	25
	Construction of physical map	31
	Rearrangement of unaligned cleavage sites	38
	DISCUSSION	43
III	CONSTRUCTION OF A SOYBEAN CHLOROPLAST GENOME CLONE BANK	50
	INTRODUCTION	50
	MATERIALS	54
	METHODS	55

	Soybean chloroplast DNA	55
	Plasmid DNA	55
	Molecular cloning of chloroplast DNA fragments	55
	Transformation and Selection	57
	Nomenclature	58
	RESULTS	58
	DISCUSSION	69
IV.	STRUCTURAL ORGANIZATION OF THE rRNA CISTRON. . .	74
	INTRODUCTION	74
	MATERIAL AND METHODS	79
	Extraction of chloroplast and plasmid DNA . .	79
	Isolation of soybean chloroplast rRNA . . .	79
	Separation of various species of soybean cp and RNA	80
	Separation of E. coli 16S and 23S rRNA . . .	80
	Restriction Digestion and Fractionation . .	81
	Recovery of labeled DNA fragments from gel .	81
	Southern and Northern Blots	81
	Radioactive labeling	82
	Hybridization	82
	Washing	83
	Autoradiography	84
	RESULTS	84
	Spatial arrangement of 16S and 23S rRNA genes	89
	Cloning	89
	Fine Mapping	90
	Mapping strategy	90
	Organization of rRNA genes	103
	Homology with E. coli and tobacco	108
	Direction of Transcription	108
	Sequence divergence studies	110
	DISCUSSION	117
	Presence of two rRNA cistrons per chloroplast DNA	117
	Organization of rDNA unit (rrn operon) . .	117
	rRNA is transcribed from 16S towards 23S . .	118
	Comparison of Soybean rRNA genes with other plants	118
	Prokaryotic feature of cpDNA	119
	Chloroplast rDNA sequence divergence . . .	120
V.	MAPPING OF CHLOROPLAST tRNA GENES.	122
	INTRODUCTION	122
	Isolation of chloroplast tRNA	125
	Two-dimensional polyacrylamide gel electrophoresis	126
	Separation and identification of individual spots	126
	Labeling of tRNA	127
	tRNA DNA Hybridization	127

RESULTS AND DISCUSSION	128
The three tRNAs have separate chloroplast genes	128
VI. POSITIONING OF PLASTIDS PROTEIN ENCODING GENES.	136
INTRODUCTION	136
Gene for Large Subunit of RUBISCO (rbcL)	137
Genes for Subunits of ATP Synthase	138
32Kd Thylakoid Protein	140
MATERIALS AND METHODS	142
Materials	142
Methods	143
M13 transformation	143
Preparation of single stranded DNA	145
In vitro conversion of single stranded DNA into Replicative Form	145
Preparation of Replicative form	146
RESULTS	146
Gene for large subunit of RUBISCO (rbcL)	146
Genes for subunits of ATPase	147
32 Kd Thylakoid protein gene (psbA)	153
Gene for cytochrome f (cytF)	153
DISCUSSION	153
BIBLIOGRAPHY	161

Chapter I
INTRODUCTION

1.1 PRESENCE OF EXTRA NUCLEAR GENETIC SYSTEM

The chloroplasts belong to a family of organelles called plastids. The presence of plastids is one of the unique and distinguishing features of the plants. The light harvesting pigment chlorophyll which imparts green color to the chloroplasts is distributed evenly in most plants but the Four O'Clock (Mirabilis jalapa) plant displays a variegated pattern i.e. green patches interspersed with pale green to white areas. Correns (1909a) observed that this variegation trait is maternally inherited suggesting the existence of genetic material (DNA) in the cytoplasm. Proof for his suggestion was obtained using the feulgen staining of Chlamydomonas chloroplast by Ris and Plaut (1962). More direct evidence was provided by Chun et al (1963) wherein they showed, using CsCl gradient, that spinach and beet DNAs resolved into a single major component and two minor components. While the major component was shown to be a nuclear derivative, the minor component was proven, in part to be chloroplastic in origin. Since then, DNA has been found in other species of plastids, namely etioplasts (Herrmann and Kowallick, 1970), chromoplasts (Herrmann, 1972) and leucoplasts (Siu et al, 1976).

1.2 PHYSICOCHEMICAL PROPERTIES OF CHLOROPLAST GENOME

In order to study the structural and functional aspects of the chloroplast genome, it is necessary to have pure, undegraded and uncontaminated chloroplast DNA (cpDNA). Several methods of chloroplast DNA isolation have been described over the years. One of the commonly used procedures was developed by Kolodner and Tewari (1975). The gravity-separated chloroplasts treated with DNAase, to degrade nuclear DNA contaminants, are lysed to release cpDNA. The lysate is extracted with phenol and precipitated with two volumes of ethanol to recover cpDNA. To obtain a cleaner preparation of DNA, CsCl - ethidium bromide density gradient centrifugation of the lysate is necessary. In this procedure the nuclear and cpDNA separate due to differential buoyant densities. Two or more CsCl - EtBr density gradients are required to obtain a reliable preparation of cpDNA. The clean, intact cpDNA molecules thus recovered, could be subjected to various analyses in order to study the physiochemical properties and probe into the genetic organization of the chloroplast genome.

The genome of any organism can be characterized with respect to complexity and organization by reassociation kinetic studies of the DNA (Britten and Kohne, 1968). The denatured DNA renatures under controlled conditions at a rate inversely proportional to its complexity (Wetmur and Davidson, 1968). The degree of reassociation can be

determined by hydroxyapatite column chromatography, spectrophotometric measurements, or by calculating resistance to S1 nuclease (Britten et al, 1974). The rapid reassociation of chloroplast DNA in comparison to nuclear DNA indicates that it is much less complex than the latter (Kung, 1977).

1.3 PHYSICAL STRUCTURE AND ORGANIZATION

With the aid of an electron microscope, Manning et al (1971) detected the presence of 40 um circular chloroplast DNA molecules in Euglena gracilis. Circularity of the chloroplast genome appears to be a constant feature of majority of the cpDNAs studied so far (Bohnert et al, 1982, Wallace, 1982, Whitfield and Bottomley, 1983) except Acetabularia which has been reported (Green, 1976) to contain linear cpDNA molecules measuring up to 200 um in length. The size of Acetabularia chloroplast genome (1.5×10^6) chloroplast genome is approximately equivalent to a medium sized bacterial DNA. The chloroplast genome of other algae, Chlamydomonas reinhardtii and Codium fragile are much smaller than Acetabularia cpDNA. (Padmanabhan and Green, 1978, Hedberg et al, 1981). The angiosperm cpDNAs (most widely studied), in contrast show drastic variations. Most of the higher plant cpDNAs are 45um long. Pea (Pisum sativum) and broad bean (Vicia faba) cpDNA are relatively smaller (39 um).

The discovery and subsequent easy availability of various restriction enzymes have facilitated studies on the structural aspects of many plant cpDNAs. Restriction enzymes are a special type of endonucleases. Depending upon the manner of cleavage, molecular weight and co-factor requirements, restriction endonucleases can be divided into two classes.

- i) Class I restriction endonucleases bind to specific nucleotide sequences but cleave at variable sites by a mechanism poorly understood. These enzymes require ATP, Mg^{2+} and Sadenosylmethionine as cofactors, and their molecular weight is approximately 300,000.
- ii) Class II restriction endonucleases recognize and cleave at specific nucleotide sequences to result in blunt or staggered ended fragments. The two DNA fragments from unrelated sources, cleaved by the same Class II restriction endonuclease can be joined by DNA ligase thereby permitting the construction of chimeric molecules. These enzymes have molecular weight in the range 20,000 - 100,000 and require only Mg^{2+} . Class II restriction enzymes have greatly facilitated the development of recombinant - DNA techniques.

Due to their ability to cleave at a specific sequence along the DNA, Class II restriction endonucleases have been

widely used to obtain restriction maps. The cleaved products of cpDNA are fractionated on agarose-EtBr gel by electrophoresis to resolve the fragments according to their relative sizes. Molecular weights of digested fragments are determined by comparing their electrophoretic mobility to a set of markers. The sum of the molecular weights of the DNA fragments produced by an enzyme is the size of the uncleaved cpDNA. The alignment of restriction products of two or more enzymes by partial, complete single, double digestion and/or by DNA/DNA hybridization yields a physical map. According to the size and genome organization, the chloroplast DNAs can be broadly separated into two categories, one comprised of plant cpDNAs ranging from 140 Kb to 180 Kb. These cpDNAs contain two copies of 20-28 Kb long region, organized in an inverted orientation with respect to each other, and referred to as inverted repeat. The second, comprised of plant cpDNAs measuring 120 to 124 Kb in contour length. These cpDNA lack an entire segment of the inverted repeat. Most of the chloroplast DNAs checked so far can be placed in the first category. Broadbean (Vicia faba) and pea (Pisum sativum) belong to the second category. Euglena sp. chloroplast is an interesting exception to the above two categories. It does not possess the inverted repeat but displays a different mode of organization. Though slight variations are noted among different strains studied, the Euglena gracilis (strains bacillus and Z) contain three tandemly repeated 5.6 Kb sequences.

The presence of inverted repeat has been confirmed by observing spontaneous renaturation of nicked, denatured circular cpDNA molecules under the electron microscope (Kolodner and Tewari, 1979). Each unit of the inverted repeat contains a single copy of 16S, 23S, and 5S rRNA genes (rDNA). The 16S and 23S rRNA sequences are separated by a spacer. Another spacer occurs between 23S rRNA and 5S rRNA genes. The spacer between 16S and 23S rDNA is larger than the spacer between 23S and 5S rDNA. In tobacco and wheat, in addition to 5S rRNA, another low molecular weight species, 4.5S rRNA, is found in association with 50S subunit (Whitfield et al, 1978). The gene for this RNA lies close to 3' end of the 23S rDNA. Sequence studies of the spacer region (Graf et al, 1980) have shown the presence of tRNA^{Ile} and tRNA^{Ala} genes in Euglena gracilis, Zea mays (corn) and Nicotiana tabacum (tobacco). However, in Spinacia oleracea (spinach), only the tRNA^{Ile}₂ gene has been detected (Bohnert et al, 1979).

Transfer RNA (tRNA) genes have been detected at other sites on the chloroplast DNA by hybridization of total 4S RNA or identified individual tRNA spots. tRNA genes have been located in Chlamydomonas reinhardtii (Malnoe and Rochaix, 1978), corn (Haff and Bogorad 1976, Mubumbila et al, 1980). Euglena spp (Hallick, 1983; El-Geweley et al, 1981). Vigna radiata (mung bean; Mubumbila et al, 1980), and spinach (Driesel et al, 1979). In spinach, tRNA genes are

encoded by the inverted repeat and the large single copy region, but in *Euglena* spp, the tRNA genes are distributed all over the map, Pillay et al (1983) have also shown the presence of tRNA in the small single copy region of soybean cpDNA.

1.4 CLONING

The development of cloning procedures have revolutionized the molecular biology studies of the chloroplast genome. Molecular cloning allows the isolation of a particular gene from the entire genome to analyze its structural and functional organization in detail. A genomic clone library is a collection of all the possible DNA sequences represented in the genome. If nuclear DNA is the starting material, a specific stretch of DNA is distributed at random among the millions of recombinant clones. However, considering the size of chloroplast genome, it is feasible to construct a clone library of relatively fewer recombinant clones, each designated to contain a specific region of the genome. The cloned cpDNA fragment can be recovered by excision of the chimeric plasmid. To date, chloroplast DNA clone banks are available for Chlamydomonas (Rochaix 1977), mung, Pisum sativa (pea) and spinach (Palmer and Thompson, 1981a), Vicia faba (broadbean ;Ko et al, 1983), Nicotiana spp (Zhu et al, 1982).

1.5 PROTEIN ENCODING GENES

In addition to coding for ribosomal and transfer RNAs, chloroplast DNA has the capacity to encode for about 100 polypeptides (Eflis, 1981) yet few protein products have been proven to be encoded by the cpDNA. Two major techniques to identify a protein-encoding sequence are as follows:

- i) In vitro transcription and translation of a special fragment. The resultant polypeptides/protein products can be identified by comparison of proteolytic digestion pattern with a known protein, two-dimensional gel electrophoresis and/or by immunoprecipitation with antibodies against chloroplast proteins. By using this technique, the genes for *rbcL* in corn (Bedbrook et al, 1979), spinach (Whitfield and Bottomley, 1980, Erion et al, 1981) and Chlamydomonas (Malnoe et al, 1979) *atpA*, *atpB* (Howe et al, 1982a) and *atpE* (Howe et al, 1982b) have been positioned.
- ii) Hybrid select translation. This method involves the hybridization of matrix immobilized total cpDNA to the cellular or chloroplastic RNA population. The set of RNAs complementary to the CpDNA retained in the columns are subsequently released by thermal denaturation. These recovered RNAs, upon translation, generate protein products partially or

completely encoded by the DNA. This technique has been utilized to position the genes for psbA (Driesel et al 1980) atpA, atpB, and atpE on spinach cpDNA (Westhoff et al, 1981).

Because the plastid genes are highly conserved, it is often convenient to use an identified gene from one plant species as a probe to position the encoding region of the same gene in other species.

1.6 PURPOSE OF THE PROJECT

The present investigation is part of a long term plan to study transcription and translation mechanisms of organellar genome in developing and aging systems. In order to examine the structural constitution, genomic organization of the soybean chloroplast DNA and to build a framework for future understanding of the expression of organellar genes during development, we decided to accomplish the following:

- i) Construct a map of the soybean chloroplast genome by positioning cleavage sites of different restriction endonucleases.
- ii) Construct a clone bank of soybean chloroplast DNA sequences in appropriate vectors.
- iii) Localize the ribosomal RNA genes, study their organization and orientation.
- iv) Map chloroplast tRNA genes on the soybean chloroplast DNA.

- v) Identify the chloroplast DNA regions encoding for plastid structural and functional proteins.

The results of these studies are mentioned in order in the following chapters.

Chapter II

PHYSICAL MAPPING OF SOYBEAN CHLOROPLAST DNA

2.1 INTRODUCTION

The molecular size (120 - 180 Kb) of the majority of cpDNAs permits the construction of restriction endonuclease cleavage maps.

The discovery and use of class II restriction endonucleases has expanded the horizons of molecular biology of the chloroplast genome. Most class II restriction enzymes employed in the construction of physical maps recognize an uninterrupted stretch of four, five or six nucleotides. The recognition site is generally a palindromic sequence, that is, it possesses twofold rotational symmetry. Hexanucleotide recognizing endonucleases cleave less frequently than endonucleases cleaving at tetranucleotides, and are therefore utilized in construction of general cleavage maps; portions of which can be fine-mapped by tetranucleotide recognizing endonuclease, if desired. Restriction endonucleases thus allow dissection of the genome at the molecular level. Mapping regions of DNA molecules is one of the many applications of the restriction enzymes, which have also facilitated the construction of recombinant plasmids, (chapter 2). The maps

of the following plant species are available to date:

Atriplex triangularis (Palmer, 1982), Atropa belladonna (Fluhr and Edelman, 1981a), Cucumis sativa (Palmer, 1982), Chlamydomonas reinhardtii (Rochaix, 1978), Euglena gracilis (Gray and Hallick, 1977), Glycine max (Singh et al, 1982, Palmer et al, 1983, Spielmann et al, 1983). Lycopersicon spp (Palmer and Zamir, 1982), Nicotiana spp (Fluhr and Edelman, 1981b, Seyer et al, 1981, Shen et al, 1982), Oenothera spp (Gordon et al, 1981a, 1981b), Osmunda cinnamomea (Palmer and Stein, 1982), Petunia spp (Bovenberg et al, 1981, Kumar et al, 1982), Phaseolus vulgaris (Palmer et al, 1983), Pisum sativum (Chu et al, 1981, Palmer and Thompson, 1981a), Sinapis alba (Link et al, 1981), Solanum spp (Palmer and Zamir, 1982), Spinacia oleracea (Driesel et al, 1979), Spirodela oligorrhiza (Van Ee et al, 1981), Triticum vulgaris (Bowman et al, 1979), Vicia faba (Koller and Delius, 1980, Ko et al, 1983), Vigna radiata (Palmer et al, 1983) and Zea mays (Bedbrook and Bogorad, 1976).

There is considerable consistency in the size and overall structural morphology of the above listed chloroplast genomes. Most of the plants contain 40-45 um circular chloroplast DNA molecules. Depending upon the developmental stage, chloroplasts of mesophyll cells may contain up to 30 copies of chloroplast DNA (Siegel, 1974). Presence or absence of inverted repeat is another variable of the above-mentioned plant cpDNAs.

In this chapter, the construction of a map of soybean chloroplast genome by positioning cleavage sites of four different endonucleases, KpnI, PvuII, SacI and XhoI, and the overall structural organization, is discussed in relation to other legume cpDNAs.

2.2 MATERIAL AND METHODS

2.2.1 Enzymes and Supplies

Restriction enzymes, normal and low melting agarose, were purchased from Bethesda Research Laboratories, Maryland. RNase and DNAase were from Sigma. The α - ^{32}P dGTP was supplied by Amersham Corporation. Plant material:

Soybean (Glycine max L) c.v Harcor seedlings were grown in plant growth chambers for 8 - 10 days under conditions of 14 hrs of light period ($36,000 \text{ erg/cm}^2 \text{ sec}$) at 25°C to allow the full expansion of the first primary leaves. The relative humidity of 80% was maintained. A day before harvesting the leaves, plants were transferred to a dark room to minimize the starch level.

2.2.2 Chloroplast isolation

Chloroplasts were prepared from kilogram quantities of soybean leaves by a modification of the procedure of Driesel et al (1979). Intact chloroplasts were obtained from soybean leaves by homogenization in a Waring Blender in a buffered medium containing 0.05 M Tris-HCl, 0.35 M sucrose

or mannitol, 0.003 M EDTA, 0.1% bovine serum albumin and 3mM β -mercaptoethanol, pH 8.0. The chloroplasts were pelleted by centrifugation at 1,000 g for 5 min. and washed once under the same conditions. The chloroplast pellet was suspended with a soft brush in 20 ml of the above buffer and centrifuged at 5,000 g for 15 min. The resulting pellet was used immediately for the preparation of chloroplast DNA or chloroplast tRNAs.

2.2.3 Isolation of DNA

DNA was prepared by a modification of the procedure of Kolodner and Tewari (1975). The chloroplast pellet (from 500g of leaves) was suspended in 20 ml of homogenization medium containing 10 mM $MgCl_2$ and 50 ug/ml DNase I to hydrolyze the nuclear DNA. After 60 min. in ice, EDTA was added to a final concentration of 50 mM and the chloroplasts pelleted at 4000 g for 10 min resuspended and washed twice in 150 mM NaCl, 100 mM EDTA, pH 8.0. The final pellet was resuspended in 10 mM Tris-HCl, 10 mM NaCl, 0.5% SDS pH 8.0, and treated with proteinase K (final conc. 50 ug/ml) for 1 to 2 hr. at 37°C. The incubation mixture was extracted with redistilled phenol saturated with 500 mM Tris-HCl, 10 mM NaCl and 0.5% SDS, pH 8.3 at room temperature. The aqueous phase was dialysed against 50 mM Tris-HCl, 10 mM EDTA and 10 mM NaCl, pH 8.3. Further treatments with proteinase K and RNaseA (final concentration 50 ug/ml) were carried out in a

dialysis bag at 37°C for 6 hr. The resulting DNA solution subjected to a final phenol treatment was concentrated by precipitation using 2.5 volumes of ethanol. For a cleaner preparation, the lysate containing cpDNA and ethidium bromide was loaded onto saturated CsCl solution in 1:1.6 ratio (El-Geweley et al, 1981) and centrifuged in a Beckman SW41 at 40,000 rpm for 10 hours at 20 °C. The lower band was collected by puncturing the side of the tube under UV light. Ethidium bromide was removed by three extractions with isobutanol saturated with NaCl and dist. water. The DNA was cleared of excessive salts by dialysis against several volumes of cold 1 mM Tris, 1 mM EDTA, pH 8.0. The DNA thus obtained, was intact and virtually free of any nuclear DNA contamination. The purity of the cpDNA was checked by agarose gel electrophoresis. DNA was stored at 4°C.

2.2.4 Cleavage of chloroplast DNA by restriction endonuclease

Up to 2 ug of DNA was digested at 37°C. with sufficient enzyme to give a complete digest. The reactions were carried out in digestion buffers specified by the supplier. In the case of double digestion, DNA was first cleaved by the enzyme requiring low strength buffer in a minimum volume (25 ul). Reactions were stopped by heating the mixture at 65 C for 5 min. The mixture was subsequently chilled at 4° C for 5 min. to separate sticky ends. Bromophenol blue (0.1% w/v) in 2% ficoll, 50 mM EDTA and 4 M urea was used as

tracking dye. Ficoll prevents the formation of U-shaped tracks (Southern 1979). The samples were quickly loaded onto 0.7% - 1.5% horizontal agarose gel slabs.

2.2.5 Gel electrophoresis

Separation and analysis of the restriction endonuclease digested DNA fragments were achieved by electrophoresis in a horizontal agarose gel slab (30 x 20 x 0.6 cm) using 0.7% - 1.5% neutral agarose gels in 100 mM Tris-acetate, 20 mM sodium acetate, 2 mM EDTA pH 8.0 for 24 - 30 hr. at a voltage of 0.5 - 3 V/cm. DNA fragments for secondary digestion were recovered from low temperature melting agarose.

2.2.6 Staining and photography

Gels stained for 30 min. in 1 ug/ml of ethidium bromide dissolved in electrophoresis buffer and destained 3- 4 times in double distilled water. The gels were photographed under long wavelength ultra violet light using a Polaroid MP-4 Land camera equipped with a yellow filter. Instant black and white Polaroid 655 and 657 films were used.

2.2.7 DNA images

The gels were processed according to Southern (1975). Two images of DNA fragments from the same gel were obtained by sandwiching the processed gel between nitrocellulose

papers as described by Smith and Summers (1980) except that 20x SSC (standard saline citrate - 150 mM NaCl and 15 mM Sodium citrate) buffer was used instead of 4X SSC.

2.2.8 Nick translation

The chloroplast DNA (4 ug) digested by a particular enzyme was dissolved in 50 mM Tris HCl pH 7.8, 5 MgCl₂, 10 mM 2 mercaptoethanol and 100 ug/ml of nuclease free BSA and labeled with α -³²P dGTP (sp. activity 3000/mmol) by the nick translation reaction according to the instruction supplied by Bethesda Research Laboratory. After the termination of labeling reaction, the labeled fragments were fractionated on 0.8% low melting agarose gel (30 x 20 x 0.5 cm). The labeled bands were cut with a sharp blade and transferred to 1.5 ml microfuge tubes. Electrophoresis buffer was added to bring agarose concentration below 0.2%. The labeled fragments were denatured by heating at 100 °C for 10 min.

2.2.9 Hybridization

Filters were pretreated with 200 ul/cm of 6X SSC, 0.5% SDS, 100 ug/ml of denatured calf thymus DNA, 5X Denhardt's (1966) for 2 - 4 hours at 68 °C in heat sealable bags (50X Denhardt's is, 5g Ficoll, 5g Polyvinyl pyrrolidone, 5g Bovine serum albumin in 500 ml of sterile H₂O). After the pretreatment, the bags were cut open at one of the corners

and the prehybridization solution squeezed out. Fresh prehybridization solution containing 10 mM EDTA and the ^{32}P labeled denatured probe DNA was added with Pasteur pipette. The bags were heat sealed using a Dazey "Seal a Meal" machine. Care was taken to release entrapped air bubbles before sealing the corner. The bags were incubated for 16 - 20 hours at 68 °C in a circulating water bath.

2.2.10 Washing

After the hybridization, the solution containing the labeled probe was recovered and used immediately for subsequent hybridizations, or discarded. The filter was carefully removed and washed four times in 2X SSC, 0.5% SDS at room temperature over an hour and twice in 0.1 X SSC, 0.5% SDS at 68 °C with gentle agitation for 90 min. each, placed on clean 3 MM Whatman paper and air dried.

2.2.11 Autoradiography

The filters were taped to 3 MM Whatman paper. The edges and sides of the filter was identified by marking with radioactive ink at appropriate locations outside the filter. The filters were exposed to Kodak XAR-5 or Fuji X-Ray "Medical" film JX (8' x 10') with Dupont intensifying screen up to 3 weeks.

2.3 RESULTS

2.3.1 Selection of endonucleases

Selection of restriction endonuclease was based upon (1) the size and sequence of the recognition site, and (2) the number of distinct fragments recovered. Table 1 shows the recognition sites (Roberts 1980) and the number of digestion fragments obtained by some restriction enzymes. Some endonucleases such as EcoR I and Hind III digest soybean cpDNA into more than 40 fragments. Four endonucleases, Kpn I, Pvu II, Sac I and Xho I were selected because they produce a simple and distinctive digestion pattern. Kpn I, Pvu II, Sac I and Xho I cleave soybean cpDNA into 16, 13, 19, and 18 fragments respectively (Table 2.1).

2.3.2 Molecular weight determination

Molecular sizes of restriction products were determined by comparing their electrophoretic mobilities with a large set of marker fragments. The marker fragments were a collection of cleavage products of DNA with EcoR I, Hind III, Sac I, Sma I, Pvu II and Xho I, and their sizes ranged from 48 Kb to 0.2Kb. Table 2.2, 2.3 and Fig 2.1 depict the molecular weights (in Kb) and stoichiometric amounts of different cleaved products of restriction enzymes Kpn I, Pvu II, Sac I, Xho I, Kpn I - Xho I, Sac I - Xho I, and Pvu II - Xho I.


The sum of individual fragments of any single digestion or double digestion yields the molecular size of the intact chloroplast genome. The average molecular size of soybean  chloroplast DNA is 151 Kb.

Table 2.1 Number of cleavage sites recognized by various restriction endonucleases on soybean chloroplast DNA.

Enzyme	Sequence Recognized	Number of Cleavage Sites
Kpn I	G G T A C [▼] C C [▲] C A T G G	16
Pvu II	C A G [▼] C T G G T C [▲] G A C	13
Sac I	G A G C T [▼] C C [▲] T C G A G	19
Xho I	C [▼] T C G A G G A C C T [▲] C	18

Recognition sequences given from 5' to 3'.

Figure 2.1 Single and double digestion products of soybean chloroplast DNA with restriction endonucleases: Xho I (lane a), Sac I - Xho I (lane b), Sac I (lane c), Pvu II - Xho I (lane d), Pvu II (lane e), Kpn I - Xho I (lane f), Kpn I (lane g), Kpn I - Sac I (lane h), Kpn I - Pvu II (lane i), Pvu II - Sac I (lane j). The molecular weights of the resulting fragments were estimated by comparing their electrophoretic mobilities against a range of DNA digests obtained separately by endonucleases EcoR I, Hind III, Sal I and Sac I. The size range is depicted in Kilobases (Kb) along the side.

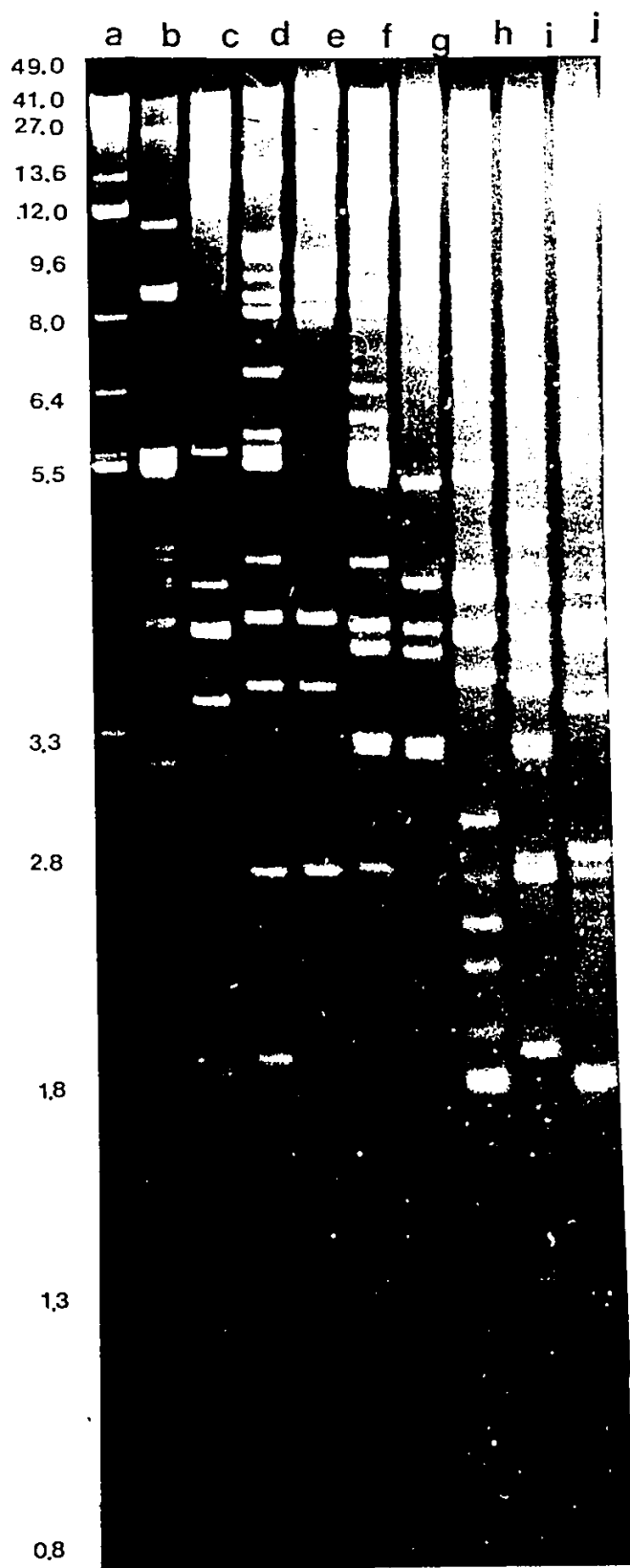


Table 2.2 Molecular sizes (in Kilobases) of soybean chloroplast DNA fragments produced by digestion with various restriction endonucleases.

Kpn I	Pvu II	Sac I	Xho I
27.5	41.0	23.2	27.0
24.0 (2X)	34.0	19.8	23.4
14.8	14.8	17.8	14.4
13.6	9.6	14.5	12.0 (2X)
12.0	9.4	13.6 (2X)	8.4
8.4	9.0	12.0	8.0
5.4	8.6	5.6	6.4
4.3	8.1	4.2	5.6 (2X)
3.9	4.1 (2X)	3.9 (2X)	5.5 (2X)
3.6	3.5	3.8	4.5
3.2	2.7	3.5 (2X)	4.0
3.1		2.8	3.3 (2X)
1.0		1.8	0.8
0.8 (2X)		1.3	0.5
150.4	151.6	152.4	150.2

2.3.3 Mapping strategy of cleavage sites

Most of the cleavage sites were ordered and arranged by DNA - DNA hybridization (Sato et al, 1977; Palmer 1982). A set of fragments obtained by Xho I digestion were individually hybridized to its own restriction products (self-hybridization) as well as to single digests of Kpn I, Pvu II and Sac I, and the secondary digests of these enzymes with Xho I. Self-hybridization may provide the information regarding the presence of repeats depending upon the constitution and orientation of the constituent fragments.

Presence of the inverted repeat: Table 2.4 shows that upon "self" hybridization, none of the Xho I probes bind to any other fragment except to itself indicating either a complete absence of a repeat or truly identical arrangement of the cleavage sites within the repeat unit, which indeed is the case. The largest Kpn fragment 27.5 Kb hybridizes to itself and 24.0 Kb Kpn fragment suggesting the presence of a repeat (Fig.2.2). Three Xho I fragments viz: 12.0 Kb, 5.5 Kb and 3.3 Kb occur in bimolar amounts and hybridize to both Kpn fragments 27.5 Kb and 24.0 Kb implying that a single copy of each Xho I fragment is present within the repeat. The orientation of the repeat was ascertained by determining the arrangement of the integrated Xho I fragments.

The 3.3 Kb Xho I probe hybridizes to 24.0 and 27.5 Kb Kpn I and to 2.7 Kb Kpn I-Xho I fragment (Table 2.5) indicating that it is flanked on the ends of Kpn I fragments

24.0 and 27.5 Kb. This Xho I probe also hybridizes to Sac I 3.5 and 3.9 Kb regions which are also present in dimeric amounts (Table 2.7). Of the remaining two fragments, 12.0 Kb binds to 3.9 Kb Sac I suggesting its proximity to 3.3 Kb region and an inverted repeat orientation of the constituent fragments, mainly 5.5, 12.0 and 3.3 Kb. Two more bimolar restricted fragments 0.8 Kb Xho I and 1.8 Kb Sac I are presumably located in part or whole within the inverted repeat. A maximum of 23.4 Kb region thus occurs as inverted repeat.

Table 2.3 Molecular sizes (in Kilobases) of soybean chloroplast DNA fragments produced by double digestion with Pvu II/Xho I and Kpn I/Xho I and Sac I/Xho I.

Pvu II/Xho I	Kpn I/Xho I	Sac I/Xho I
14.6	24.0	19.8
14.4	13.2	11.0
10.2 (2X)	12.0 (2X)	8.7 (2X)
9.6	8.6	8.5
9.0	6.2	5.6 (4X)
8.4	5.6	5.5 (2X)
8.0	5.5 (2X)	5.4
	5.4 (2X)	4.6
6.0	4.5	4.5
5.6	3.9 (2X)	4.2
5.5 (2X)	3.6 (2X)	3.9
4.5	3.2	3.8
4.2	3.15	3.2
3.9	3.1	3.1 (2X)
3.8	2.7 (2X)	2.8 (2X)
3.6	2.4	2.6 (2X)
2.7 (2X)	2.3	2.4
2.4	1.8	1.8 (3X)
2.0 (3X)	1.2	1.3
0.8 (2X)	1.0	0.8 (4X)
0.7 (2X)	0.8 (2X)	0.5
0.6 (3X)	0.7 (2X)	
	0.5	
152.6	150.1	149.6

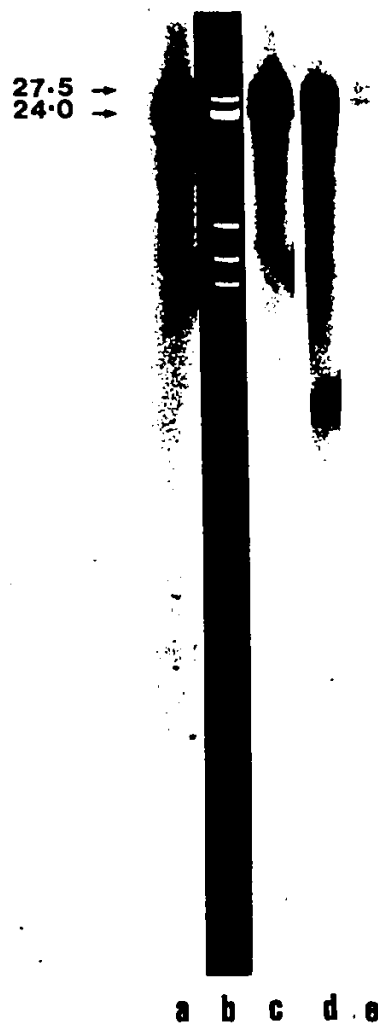


Figure 2.2 Hybridization of α - 32 P dGTP labeled 27.5 Kb Kpn I fragment and three Xho I fragments, 12.0 Kb, 5.5 Kb and 3.3 Kb to nitrocellulose strip containing Kpn I digests. Lane (a) binding of 27.5 Kb Kpn I probe to itself and 24.0 Kb Kpn I region; lane (b) ethidium bromide stained pattern of Kpn I cleavage products; lanes (c, d and e) binding pattern of 12.0 Kb, 5.5 Kb and 3.3 Kb Xho I fragments respectively.

Table 2.4 Summary of hybridization of (α - 32 P) dGTP labeled Xho I fragments with its own restriction pattern.

(α - 32 P) dGTP Xho I Probe (Kb)	Single digest Xho I fragments (Kb)
27.0	27.0
23.4	23.4
14.4	14.4
12.0 (2X)	12.0
8.4	8.4
8.0	8.0
6.4	6.4
5.6 (2X)	5.6
5.5	5.5
4.5	4.5
4.0	4.0
3.3 (2X)	3.3

Sizes of the fragments are in Kb.

Table 2.5 Summary of hybridization of (α - 32 P) dGTP labeled Xho I fragments with single digest of Kpn I, Pvu II and Sac I and double digests of Xho I/Kpn I, Xho I/Pvu II and Xho I/Sac I.

(α - 32 P) dGTP probe Xho I	Single digest Kpn I fragments	Double digest Xho I/Kpn I fragments
27.0	14.8, 8.4, 5.4, . 3.2, 3.1, 1.0	8.6; 5.4 (2X)*, 3.2, 3.1, 1.0
23.4	24.0	23.4
14.4	13.6, 4.3	13.2, 1.2
12.0 (2X)	27.5, 24.0	12 (2X)*
8.4	24.0, 3.9	3.9, 3.6, 0.4 ^a
8.0	27.5, 12.0, 3.6	3.6, 2.3, 1.8
6.4	14.8	6.2, 0.2 ^a
5.6 (2X)	12.0, 8.4, 4.3	12.0, 3.2, 2.4
5.5 (2X)	27.5, 24.0	5.5 (2X)*
4.5	27.5	4.5
4.0	12.0	4.0
3.3 (2X)	27.5, 24.0, 0.8	2.7 (2X)*

^a Deduced theoretically

* Hybridization with both fragments

Sizes of fragments are in Kb.

2.3.4 Construction of physical map

The entire physical map can be organized by aligning the Kpn I, Sac I and Pvu II sites on the Xho I probes. (Tables 2.5, 2.6, 2.7) The principle of arranging discrete order of restriction fragment is illustrated by positioning Kpn I, Sac I and Pvu II sites on 14.4 Kb Xho I. This probe hybridized to 13.6 and 4.3 Kb Kpn I single digests (Table 2.5) thereby indicating that 14.4 Kb overlaps these Kpn I fragments. The extent of the overlap was determined from the hybridization pattern of Kpn I - Xho I double digests (Table 2.5). It revealed that 14.4 Kb Xho I hybridized to 13.2 and 1.2 Kb Kpn I - Xho I double digests. Note that the sum of double digest fragments is equal to the probe fragment, 14.4 Kb Xho I. This delineated the extent of 14.4 Kb over the two Kpn I fragments; 13.2 Kb of 14.4 Kb was stretched over 13.6 Kb Kpn I and the remaining 1.2 Kb was flanked over 4.3 Kb Kpn I. Similarly, Table 2.6 shows that 14.4 Kb is completely contained in 41.0 Kb Pvu II fragment because the probe hybridized to fragment equivalent to its own size in PvuII - Xho I double digest. Accordingly, 14.4 Kb Xho I was found to overlap 14.5, 12.0 Kb Sac I (Table 2.7 and Figs. 2.3, 2.4) In similar fashion, the sites of Kpn I, Pvu II and Sac I were positioned on all the Xho I probes. The discrete order of the restriction sites established the circularity of the genome as the cleavage sites of Kpn I, Pvu II, Sac I and Xho I could only be organized in a circular array (Fig. 2.7).

Table 2.6 Summary of hybridization of (α - 32 P) dGTP labeled Xho I fragments with single digest of Kpn I, Pvu II and Sac I and double digests of Xho I/Kpn I, Xho I/Pvu II and Xho I/Sac I.

(α - 32 P) dGTP probe Xho I	Single Digest Pvu II fragments	Double Digest Xho I/Pvu II fragments.
27.0	14.6, 9.4, 8.6	14.6, 8.6, 7.0
23.4	9.6, 9.0, 3.6	9.6, 9.0, 3.5, 4.1
14.4	41.0	14.4
12.0 (2X)	41.0, 34.0, 2.7	10.2 (2X)*, 2.0 (2X)**
8.4	41.0	8.4
8.0	34.0	8.0
6.4	9.4, 8.1	4.2, 2.4
5.6 (2X)	41.0, 34.0, 8.6	5.6, 3.6, 2.0
5.5 (2X)	41.0, 34.0	5.5 (2X)*
4.5	34.0	4.5
4.0	8.1	3.9
3.3 (2X)	4.1, 2.7	2.7, 0.6 ^a

* Hybridization observed with both fragments

a Deduced theoretically

Sizes of fragments are in Kb.

Table 2.7 Summary of Hybridization of
 $(\alpha\text{-}^{32}\text{P})$ dGTP labeled Xho I fragments
 with single digest of Kpn I, Pvu II
 and Xho I/Sac I

$(\alpha\text{-}^{32}\text{P})$ dGTP probe Xho I	Single Digest Sac I fragments	Double Digest Sac I/Xho I fragments
27.0	13.6, 5.6, 4.2, 2.5, 1.8, 1.3	11.0, 5.6, 4.2, 2.5, 1.8, 1.3
23.4	19.8, 1.8 (2X)*	19.8, 1.8 (2X)*
14.4	14.5, 12.0	8.5, 5.6
12.0 (2X)	23.2, 17.8, 3.9 (2X)*	8.7 (2X)*, 3.1 (2X)*
8.4	17.8, 14.5	5.6, 2.8
8.0	23.2, 13.6	4.6, 3.2
6.4	13.6, 3.8	3.8, 2.5
5.6 (2X)	13.6, 12.0	5.6 (2X)*
5.5 (2X)	23.2, 17.8	5.5 (2X)*
4.5	23.2	4.5
4.0	13.6	4.0
3.3 (2X)	3.8, 3.5	2.6, 0.8

*Hybridization observed with both fragments
 Sizes of fragments are in Kb.

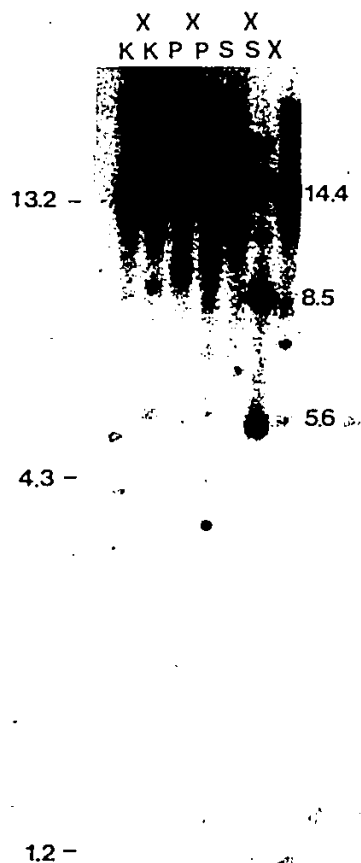


Figure 2.3 Hybridization of Nick translated 14.4 Kb Xho I fragment to single digests of Kpn I (K), Pvu II (P), Sac I (S), Xho I (X) and double digest of Kpn I - Xho I (XK), Pvu II - Xho I (XP), Sac I - Xho I (XS). Size (in Kilobase) of some of the fragments hybridizing to the probe are mentioned along the sides.

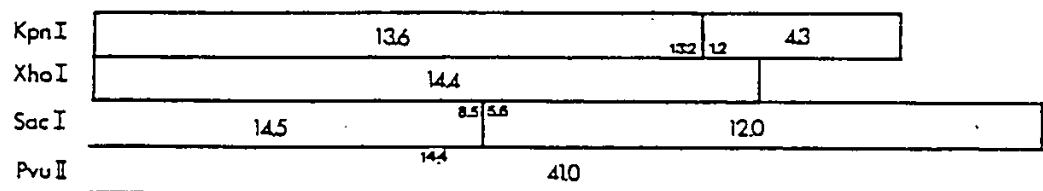


Figure 2.4 Positioning of Kpn I, Pvu II and Sac I cleavage sites on Xho I fragment from hybridization data by measuring the extent of overlap. The size of the overlap is mentioned in smaller print.

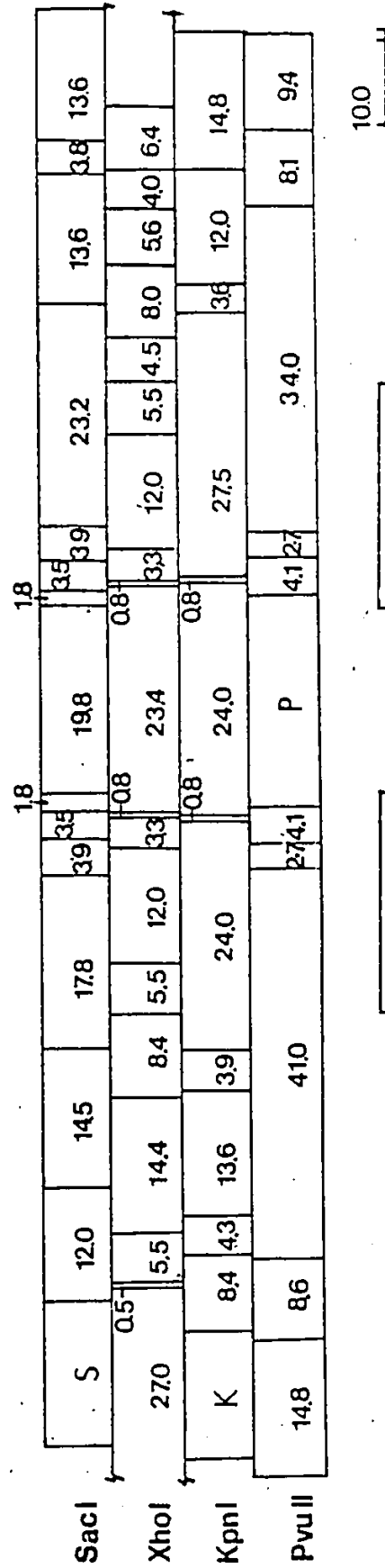


Figure 2.5 Linear map of soybean chloroplast DNA showing the order of various restriction sites of Sac I, Xho I, Kpn I and Pvu II. The 23.5 Kb inverted repeat region is represented by the open bars. K, P and S are the three regions containing cleavage sites, left unaligned by Xho I hybridization.

2.3.5 Rearrangement of unaligned cleavage sites

It is difficult to assess the precise organization of cleavage sites of any given enzyme, by hybridization data alone, if they are fully contained within the probe fragment. To illustrate, the 27.0 Kb Xho I region hybridized to 13.6, 5.6, 4.2, 2.8, 1.8 and 1.3 Kb Sac I fragments (Table 2.7 and Fig. 2.6). Only those Sac I cleavage products can be unambiguously positioned which are spanned on the edge of 27.0 Kb Xho I by determining the overlapping fragments. Fig. 2.6 and Table 2.7 indicates that a large portion (11.0 Kb) of ~~13.6~~ 11.0 Kb Sac I is stretched across 27.0 Kb Xho I cleavage product, whereas the rest of the five Sac I fragments are present within the remaining portion. Four Kpn fragments viz: 5.4, 3.2, 3.1 and 1.0 Kb and three Pvu II fragments viz: 9.6, 9.0 and 3.5 Kb were similarly left unarranged by hybridization. The actual order of unaligned Sac I segments were determined by hybridizing $\alpha^{32}\text{P}$ -Kpn I translated Kpn I fragments against Sac I digestion pattern. From the hybridization results (Table 2.8) it was clear that the arrangement of Kpn I fragments was 5.4, 3.2, 3.1 and 1.0 Kb, and the Sac I fragments were ordered as: 5.6, 4.2, 2.8, 1.3 and 1.8 Kb. The inability to completely separate the nearly same 3.2 and 3.1 Kb Kpn I probes on 1.0% low melting agarose consequently resulted in the observed cross-contamination. The Pvu II fragment pattern was established by comparison of Pvu II and Sac I

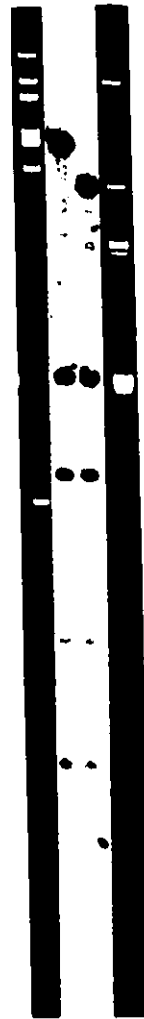
single digest with Pvu II - Sac I double digest (Fig. 2.1). Note that both 9.6 and 9.0 Kb Pvu II are absent in Pvu II - Sac I double digest whereas 3.5 Kb is retained indicating the sequence 9.6 - 3.5 - 9.0. Because these three Pvu II fragments encompass the single copy region and are delimited by the inverted repeat, it is therefore difficult to ascertain their exact position with respect to fragments lying outside the inverted repeat.(Fig. 2.7).

Table 2.8 Positioning of unaligned Kpn I and
Sac I fragments;
Hybridization of α - ^{32}P dGTP labeled
Kpn I fragments against Sac I
digestion pattern.

Radioactive Kpn I fragments	Filter bound Sac I fragments
5.4	5.6, 4.2, 2.8
3.2	13.6, 2.8, 1.8, 1.3
3.1	13.6, 2.8, 1.8, 1.3
1.0	13.6

Sizes of the fragments are in Kb.

27.0 Xho I



s a b xs

Figure 2.6 Hybridization of Nick translated 27.0 Kb Xho I fragment to the single digest of Sac I (lane a) and double digest of Xho I - Sac I (lane b). Digestion patterns of Sac I (s) and Xho I - Sac I (XS) are depicted in the flanking lanes.

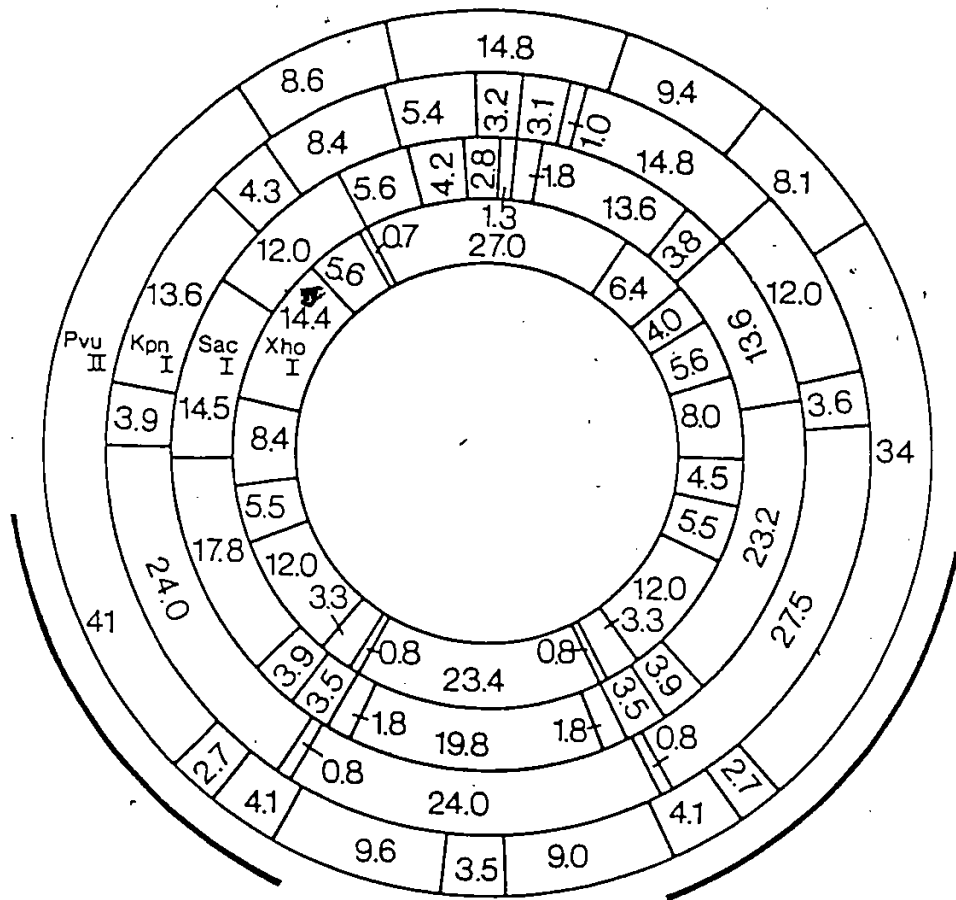


Figure 2.7 Four-track circular map of soybean chloroplast DNA depicting the discrete order of all the digestion products of Kpn I, Sac I, Pvu II and Xho I. The thick arcs outside the map represent the presence of the inverted repeat. Sizes of the fragments are in Kb.

2.4 DISCUSSION

Previous studies from this laboratory (Pillay and Gowda, 1981) have established degradative changes in chloroplastic tRNAs and tRNA synthetases in aging soybean leaves/cotyledons. In order to understand the tRNA gene expression and function in chloroplasts, it was necessary to first study the structural and functional organization of the soybean chloroplast genome. Physical mapping is the foremost step to any such examination because it lends a framework to position and investigate different genes.

The structural organization of the soybean chloroplast DNA was mapped by cleavage sites of four restriction endonucleases; Kpn I, Pvu II, Sac I and Xho I. The objective of the mapping technique is to determine the number and extent of overlaps between restriction fragments. The procedure adopted requires a minimum of two sets of restriction digests to construct a physical map. The major approaches used in mapping restriction sites are:

- i) Reciprocal Digestion; The digestion products of an enzyme are eluted out of the gel and redigested with the second enzyme. Similarly, the digestion products of the second enzyme are eluted out of the gel and redigested with the first enzyme. The extent of overlap i.e. the region of similarity between two fragments, is established by comparing the common subfragments obtained during reciprocal

digestion of the individual fragments (Herrmann et al, 1980, Link et al, 1981, Gordon et al, 1981).

- ii) DNA-DNA Hybridization; The overlaps are determined by hybridizing one set of labeled restriction fragments against another group of filter-bound restriction fragments. The hybridization method of defining the discrete order of restriction fragments is based upon complementarity. A number of rapid variations of these methods have been developed to facilitate mapping. One of these approaches (Smith and Brinsteel, 1976) is an elegant procedure to position cleavage sites for one or more restriction enzymes concomitantly. The rationale behind this method is the ability to resolve all the partial digests of a given enzyme. The DNA fragment to be mapped is labeled at the 5' end by T4 DNA kinase and γ -P-ATP, partially digested with the given endonuclease and fractionated. The longest labeled partial digest is eluted out of the gel and again partially digested by the same or another enzyme. The labeled incomplete digests are fractionated on the agarose and their position determined by autoradiography of the gel. The sequence of the restriction sites can be read in order from the autoradiograph.

The success of the Smith and Brinstiel (1976) method of mapping depends upon the resolution of all the possible partial fragments. Gathering the whole spectrum of well resolved incomplete digests of cpDNA with a particular enzyme on 0.8% agarose is rather tedious. This method is well suited for mapping small fragments (25 Kb). Reciprocal digestion was not selected because it requires large amounts of cpDNA, and depends upon the ease of purifying individual fragments out of the gel.

— Yet another refined technique to map cleavage sites has been described by Sato, et al (1977). In this method, the digestion products of a restriction enzyme are fractionated on an agarose gel (such that the entire width of the gel is covered with the loaded sample) denatured and transferred to the nitrocellulose filters (Southern, 1975). DNA fragments generated by digestion of P-labeled DNA by the second enzyme are similarly fractionated and transferred to the same nitrocellulose paper at right angles to the first transfer. The transfer is carried out under conditions which enable the second set of labeled fragments to hybridize with originally fixed DNA fragments.

The techniques described here are the ones commonly used to position restriction sites. The choice of the method depends upon the needs of the worker. Because DNA-DNA hybridization and two-dimensional hybridization methods Sato et al (1977) are based upon complementarity, the unambiguous

positioning of restriction sites is very low. DNA-DNA hybridization was selected over two-dimensional hybridization technique because it is relatively inexpensive and time-saving.

In our studies labeled Xho I restriction fragments were used as hybridization probes against a collection of nitrocellulose fixed single digests of Kpn I, Pvu II, Sac I and Xho I restriction enzymes plus double digests of Kpn I - Xho I, Pvu II - Xho I, and Sac I - Xho I.

Hybridization of such probes to different single digests will determine the overlaps whereas hybridization to double digests will define the extent of the overlap, hence precisely positioning restriction sites. Hybridization of labeled Xho I probes to filterbound Xho I cold fragments (self hybridization) will detect the regions of homology or repeated zones.

The estimated size of soybean cpDNA (151 Kb), corresponds to the sizes of other plant chloroplast genomes; spinach (Driesel et al, 1979), Oenothera sp (Bovenberg et al 1981), Atriplex (Palmer, 1982). Another feature of the soybean chloroplast genome that is common to the other plants is the presence of the classical inverted repeats which places it with first category of known plastid genome organization.

The soybean chloroplast DNA is a circular molecule which is organized into two unequal single copy regions of

84 Kb (large) and 20 Kb (small). The large and small single copy regions are interrupted by a single 23.5 Kb repeat region on either side. The repeated region is organized in an inverted orientation. Spielmann et al (1983) have also reported the construction of physical map of soybean chloroplast genome. Their results are comparable to our results. According to their estimate, the size of soybean cpDNA is between 150 - 152 Kb and length of inverted repeat is about 20 - 23 Kb; In their opinion the small single copy region is 24 Kb, nearly 4 Kb bigger than our calculations (Palmer et al ,1983). In most of the higher plant cpDNAs, the length of each unit of the inverted repeat is between 22 and 24 Kb though cucumber (Cucumis sativa) possesses a much shorter (14 Kb) inverted repeat (Palmer, 1982).

The inverted repeats found in a variety of cp genomes studied has been of considerable interest to plant molecular biologists. This region contains genes for ribosomal and transfer RNAs. The presence of the inverted repeat has been attributed to a more stable configuration (Kolodner and Tewari, 1979, Fluhr and Edelman, 1981a Palmer and Thompson, 1981a).

Work is in progress in several laboratories to assess the homologies ~~between~~ various chloroplast genomes and their relationship, ~~in~~ any to the presence of an inverted repeat.

Palmer et al, (1983) compared the overall structure and nucleotide sequence similarities of three legume

chloroplast DNAs viz. soybean, common bean and mung bean by heterologous filter hybridizations. They showed that the common bean and mung bean cpDNAs are essentially colinear and differ by 5 - 6% base sequence divergence. On the other hand, mung bean and soybean, although show a similar arrangement of restriction sites, two regions of deletions/additions, each totaling almost 5 Kb in size, have been recognized at the ends of the large single copy DNA region. The mung bean and soybean cpDNAs indicate a greater (10-13%) base sequence divergence. Base substitutions are distributed at random in the chloroplast DNAs.

To find if the linear order of restriction sites is conserved between chloroplast DNAs from various plants, Palmer and Thompson (1982) compared the structural and large sequence elements by heterologous hybridizations. They found more rearrangements in the structural organization of chloroplast genome (broad bean, pea) lacking the inverted repeat as compared to chloroplast DNAs containing it (spinach, petunia, cucumber). Sweet pea (Lathyrus odoratus) and alfalfa (Medicago alfalfa) are two other plants which seem to have deleted an entire segment of the inverted repeat (and depict more inversions/deletions in the gross structural organization in comparison to the spinach group (Palmer, personal communication). All the plants lacking inverted repeat are members of the subfamily Papilionoideae (Fernald, 1950) of family Leguminosae. The absence of the

inverted repeat apparently is a feature common to a small group of plants of the same subfamily whereas presence of the inverted repeat is the general mode of structural organization of plant cpDNAs. A number of workers (Bedbrook et al, 1977, Kolodner and Tewari, 1979, Fluhr and Edelman, 1981a, Palmer and Thompson, 1982) have repeatedly mentioned that the occurrence of the inverted repeat as compared to loss of an entire segment in a few plants belonging to a particular sub group, is a consequence of evolutionary change.

Chapter III
CONSTRUCTION OF A SOYBEAN CHLOROPLAST GENOME
CLONE BANK

3.1 INTRODUCTION

A longstanding goal of molecular biologists is to understand the organization and expression of genes. Until recently, the few conventional methods available for purification of a particular gene were relatively tedious, inefficient and restricted in their approach. However, rapid advances in cloning techniques and recombinant DNA technology have made the isolation of specific genes and their gene products possible. Several gene products, beneficial to scientific research, medicine or industry, have been isolated by genetic engineering and molecular cloning. Synthesis of insulin (Ullrich et al, 1977) somatomammotropin (Shine et al, 1977) and interferon (Taniguchi et al, 1980a) by recombinant technology are some of the well established examples.

Cloning is, in principle, a two-step procedure. The first step involves the joining of a given DNA segment with a suitable vector to form a stable recombinant molecule. The second involves the uptake and subsequent propagation of this recombinant plasmid by an appropriate and compatible

host. The complete gene library (gene bank) is a representation of the entire population of DNA sequences. Genomic DNA clone library is more comprehensive than the cDNA (complementary DNA) clone library because the latter just represents the abundant mRNA population and does not contain the intervening sequences (introns) or the flanking regions.

Construction of a genomic library requires the collection of insert DNA fragments representing the total genome. The DNA fragments, appropriate for cloning, can be obtained by mechanical shearing (Wensink et al, 1974) or by partial digestion with one or more restriction endonucleases (Sinsheimer, 1977). On the other hand the preparation of a cDNA library requires the complementary DNA synthesized from the purified mRNA fractions. Next step in the construction of either type of genomic or cDNA libraries is ligation of DNA fragments with a suitable vector in the presence of DNA ligase under conditions which will result in circular, monomeric recombinant DNA molecules. Several vector systems are available to suit the needs of the worker. Plasmids, lambda (λ) phage and cosmids are the major vector systems used to clone and propagate foreign DNA sequences in Escherichia coli (E.coli).

Plasmids are extra chromosomal genetic elements of many bacterial systems and are associated with sex factors, antibiotics resistance, production of colicins and

enterotoxins. In nature, bacterial plasmids range in size from few kilobases to over 200 Kb but plasmids utilized for cloning are generally less than 8 Kb. Naturally occurring plasmids can be modified by recombinant techniques to result in a suitable cloning vehicle in bacteria. A plasmid used as cloning vector should possess several features which include (a) small size, (b) capability of autonomous replication, (c) presence of identifiable marker with single cleavage sites of many restriction endonucleases, restriction sites outside the markers, should be positioned such that cloning at these sites should not alter the replication properties of the plasmid. (d) ability to be transformed at high frequency. pBR322 (Bolivar et al, 1977) is an example of versatile cloning vector containing the above mentioned properties.

Bacteriophage λ vectors: The success of λ phages to be adapted as cloning vehicle essentially depends upon their ability to carry out the lytic (infection) cycle without the participation of the central one-third of their genome. Two types of λ vectors have been created; insertion and replacement vectors. As the name implies, in insertion vectors, foreign DNA sequences could be introduced at the insertion sites whereas in replacement vectors the non-essential portion of the λ genome could be substituted by foreign DNA. Large DNA fragments, up to 25 Kb can be inserted in λ vectors. More information regarding the

construction of bacteriophage λ vectors can be obtained from William and Blattner (1980), Dahl et al (1981) and Maniatis et al (1982).

Cosmids are those modified plasmids which contain phage cohesive (cos) ends (Collins and Brunning, 1978). These cloning vehicles were specially designed (Collins and Hohn, 1979) to clone large (35 - 50 Kb) fragments. An insert fragment ligated with cosmid can be packaged by the phage packaging system in vitro and transduced into E. coli at high frequency, (Feiss et al, 1977). Inside the bacterial cell, due to the absence of the required lambda genes to carry out lysogeny or lytic cycle, cosmid can not perform as infectious lambda particle. Instead, it replicates and behaves as a plasmid. Despite the advantage in packaging large fragments, cosmids haven't been widely used in constructing plant genome libraries.

The resulting recombinant DNA molecules are introduced into bacterial host (generally E. coli) by transformation or transduction. The last step in constructing a clone bank is the selection of bacterial colonies in case of cosmid or plasmid cloning or plaques in case of lambda bacteriophage cloning, containing the desired foreign DNA sequences. This generally involves nucleic acid hybridization (Grunstein and Hogness, 1975).

In this chapter, the methods describing the cloning of soybean chloroplast DNA fragments in plasmids, pBR322 and

pDPL13 are discussed along with some of the practical problems encountered during pHC79 cosmid cloning.

3.2 MATERIALS

Enzymes: Restriction enzymes were purchased from Bethesda Research Laboratory, Maryland. Calf intestine alkaline phosphatase was obtained from Boehringer, Mannheim; and T4 DNA Ligase was from New England Biolabs, Boston.

Bacterial Strains: E. coli strain HB101 (pro⁻ leu⁻ thi⁻ Lac/Y⁻ hsdR⁻ endA⁻ recA⁻ rpsL20 ara-14 gal k2 xyl-5 mtl-1 Sup E44) was from The American Type Culture Collection, Washington D.C., U.S.A.

Media.

LB (Luria-Bertani) Medium: This medium consisted of 10 g Bactotryptone, 5 g Bacto-yeast extract and 10 g NaCl per litre of solution.

M9 Medium: The constituents of this medium were as follows: 2 g Casamino acids, 10 ml of 100 mM MgSO₄, 1 ml of 100 mM CaCl₂ and 870 ml of distilled water. After autoclaving and cooling, 100 ml of sterile 10 X salt solution containing 70 g Na₂HPO₄, 30 g KH₂PO₄, 10 g NH₄Cl, 5 g NaCl per litre of solution, 20 ml of 20% sterile glucose were added.

To make solid medium for plates, 13.0 g of Bacto-Agar (Difco) was added to the above medium before autoclaving. Appropriate antibiotics were added after the medium was

cooled. The concentration of antibiotics, ampicillin and tetracycline, used for routine selection were 50 ug/ml and 25 ug/ml respectively. (35 - 50 ug/ml) and Tetracycline (25 ug/ml) respectively.

3.3 METHODS

3.3.1 Soybean chloroplast DNA

The chloroplast DNA was prepared as described in Chapter one.

3.3.2 Plasmid DNA

Plasmid DNA was prepared by SDS-lysis of bacterial strains harbouring plasmids, pBR322 and pDPL 13 according to Godson and Vapnek (1973).

3.3.3 Molecular cloning of chloroplast DNA fragments

Restriction digestion and electrophoresis digestions were carried out according to suppliers' instructions. The DNA was fractionated on 0.8 - 1.2% agarose gel at 1 - 2 v/cm for 10 - 25 hours at room temperature. The gels were stained in 0.5 ug/ml of ethidium bromide in double distilled water for 15 min. and destained by immersing in double distilled water for 30 min. and the DNA visualized under long UV light.

Processing of cpDNA Restriction Fragment: Pvu II or Sac I of cpDNA fragments were extracted twice with phenol-

chloroform mixture and twice again with chloroform. The aqueous phase was cleared of chloroform by extraction with ether. Traces of ether were removed by heating the contents at 60 to 65 °C for 10 min. Sodium acetate, pH 5.2, to a final concentration of 0.25 M and exactly 2 volumes of chilled absolute ethanol were added to the solution containing DNA fragments. The contents were stored at -20 °C for 2 - 3 hours and cpDNA fragments were recovered by centrifugation, and traces of ethanol cleared by cotton swab. The cpDNA fragments were suspended in double distilled water to obtain a concentration of 1 mg/ml.

Dephosphorylation of plasmid DNA: Plasmid DNA was dephosphorylated by calf intestine alkaline phosphatase at 37 °C for an hour in order to avoid recircularization of the plasmid.

Ligation: Ligation of dephosphorylated plasmid DNA and insert DNA was achieved by T4DNA ligase according to the reaction conditions specified by the commercial suppliers. To ensure proper ligations, conditions should be such that they initially favour the joining of one end of the fragment to the plasmid and aid circularization of the hybrid molecules as the reaction proceeds (Dugaiczyk et al 1975). The plasmid vector: DNA fragment concentration (1 ug/ul) of 1:1, 1:3 in the ligation reaction generates enough chimeric plasmid to score $10^4 - 10^5$ recombinant colonies / ug of DNA. Blunt-ended ligation fragments usually required 50 to 80 times more T4 DNA ligase than the cohesive end ligation.

3.3.4 Transformation and selection

Bacterial strain HB101 (Boyer and Roulland Dussoix, 1969) of E. coli was transformed by recombinant plasmids according to the procedures of Mandel and Higa (1970). Approximately 50 ml of LB medium was inoculated with 0.5 ml of an overnight HB101 culture. The growth of the bacterial cells was terminated at O.D.₆₀₀ = 0.6 by chilling the cells at 4°C. The ability to take up DNA is highest for bacteria in mid logarithmic phase. About 10 ml of the bacterial culture were withdrawn, centrifuged at 4000 g for 5 min. at 4°C. The bacterial pellet was immediately chilled in icy water (2 - 4°C). All the succeeding steps unless mentioned otherwise were performed at 4°C.

The bacterial pellet was suspended in 5 ml of pre-chilled 50 mM CaCl₂ for 20 minutes. The suspension was recentrifuged at 4000 g for 5 min. and the pellet resuspended in 1 ml of ice cold 50 mM CaCl₂. About 200 ul of cells were added to the ligation solution containing chimeric plasmids and stored on ice for 20 min. Transformed bacterial cells were subjected to heat shock at 42°C for 2 min. before enriching them in LB broth for 2 hours at 37°C. Up to 200 ul of bacterial cells were plated on M9 medium supplemented with Ampicillin (50 ug /ml). The plates were incubated at 37°C for 10 - 12 hours. Ampicillin resistant (Amp) colonies thus recovered were used to analyze the recombinant DNA plasmids.

Rapid isolation of plasmid DNA: For quick screening of the inserts, rapid isolation of plasmid DNAs from small cultures by SDS lysis (Birnboim and Doly 1979) or by boiling (Holmes and Quigley, 1981) was carried out.

Storage of bacteria/recombinant clones: Bacterial strains with or without recombinant plasmids were stored in 53% glycerol (BDH). Five ml of M9 medium was inoculated with selected bacterial colonies and incubated overnight at 37 . The cells were harvested and suspended in 10 mM sterile MgSO₄ and glycerol to a final concentration of 53%. The colonies could be safely stored for about a year.

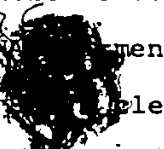

3.3.5 Nomenclature

The nomenclature of the recombinant plasmid is based upon: 1) The first letter of the vector (pBR322 or pDPL13) used; 2) The first letter of the insertion site employed for cloning (Pvu II or Sacc I), and 3). The numerical order of the inserted fragments (To illustrate, 14.8 Kb Pvu II is the third largest Pvu II fragment and is inserted in pBR322, hence it will be pBP3).

3.4 RESULTS

The initial strategy was to clone Pst I digestion products of soybean cpDNA in the plasmid, pBR322. Pst I was selected because it generated nine distinct fragments ranging in size from 0.9 Kb to 41.0 Kb; and a suitable

vector, pBR322 containing a single Pst I site in the ampicillin gene was available. All except three Pst I fragments, 41.0, 31.0 and 30.0 could be cloned in the chosen vector and were indeed inserted at Pst I site in the ampicillin drug marker (unpublished results) by insertional inactivation. Insertional inactivation is the loss of activity or resistance upon introduction of a DNA segment at insertion site contained within an identifiable marker. Attempts to package three large Pst I fragments in cosmid pHC79 were unsuccessful. It was observed that cpDNA fragments did ligate to dephosphorylated pHC79 to form DNA molecule, but the molecules could not be transduced into E.coli at high frequency (4×10^5 hybrid clones/ μ g of foreign DNA) described by Collins (1979). Instead, the few colonies recovered were the ones containing cosmids lacking inserts. Perhaps the recombinant cosmid with cpDNA inserts, somehow are either not packaged, or else cannot replicate inside bacterial cells upon transduction. Therefore the idea of constructing Pst I-based clone bank was abandoned and a new approach to raise the clone library was considered.

The new strategy involved cloning of Pvu II and Sac I cpDNA fragments in suitable vectors. Pvu II generated 13 blunt-ended soybean cpDNA fragments ranging in size from 7 Kb to 41.0 Kb, whereas  cleaved the soybean chloroplast genome into 19 fragments with cohesive ends 

ranging from 1.3 Kb to 23.2 Kb. The entire collection of Pvu II (41.0, 34.0, 14.8, 9.6, 9.4, 9.0, 8.6, 8.1, 4.1(2x), 3.5, and 2.7(2x) Kb) except 41.0 and 34.0 and Sac I (23.2, 19.8, 17.8, 14.5, 13.6, 12.0, 5.6, 4.2, 3.9(2x), 3.8, 1.8(3x), and 1.3Kb) fragments. were below the upper limit of plasmid cloning.

pBR322 was cut open by Pvu II, dephosphorylated with calf thymus phosphatase to prevent recircularization of the vector DNA, and ligated to Pvu II digestion products in the presence of T4 DNA ligase. The resulting recombinant plasmids were introduced in CaCl₂ treated bacterial cells and selected on ampicillin plates. All the recombinants thus recovered were individually analyzed by the rapid isolation procedure of Birboim and Doly (1979).

The collection of Pvu II fragments cloned in pBR322 are shown in Fig.3.1 and Table 3.1. Notice that 2.7 Kb and 8.1 Kb Pvu II fragments are missing from the set of Pvu II recombinant clones. Not even a single recombinant clone containing the cpDNA insert was recovered from over 1500 colonies screened by quick plasmid isolation procedures. Apparently the size of the insert may not be the only criteria for successful cloning; perhaps sequence of the insert may play a role in stable and efficient cloning.

Sac I fragment overlapping or covering the uncloned regions of Pvu II, were inserted at Sac I site in pDPL13, a derivative of pBR322. pDPL13 houses a 79 bp polylinker

containing cleavage site for Sac I and twelve other enzymes (Gendel et al 1983). The Sac I recombinant clones are displayed in Table 3.2. Figure 3.3 depicts the collection of eight Sac I fragments. Though the insertion sites in both vectors are outside their respective drug markers, the recovery of recombinant clones was not a formidable task because the transformation and insertion frequencies were high. Insertion rates of 70 - 80% were generally obtained with fragments up to 15 Kb. The frequency of cloning beyond 15 Kb decreased with increasing size of the fragments.

In order to clone the largest 23.2 Kb Sac I fragment, the plasmid pDPL13 was dephosphorylated and ligated with agarose eluted 23.2 Kb Sac I in 1:1 ratio (1 µg/ul). Of the 34 colonies scored, five contained the insert (Fig. 3.4).

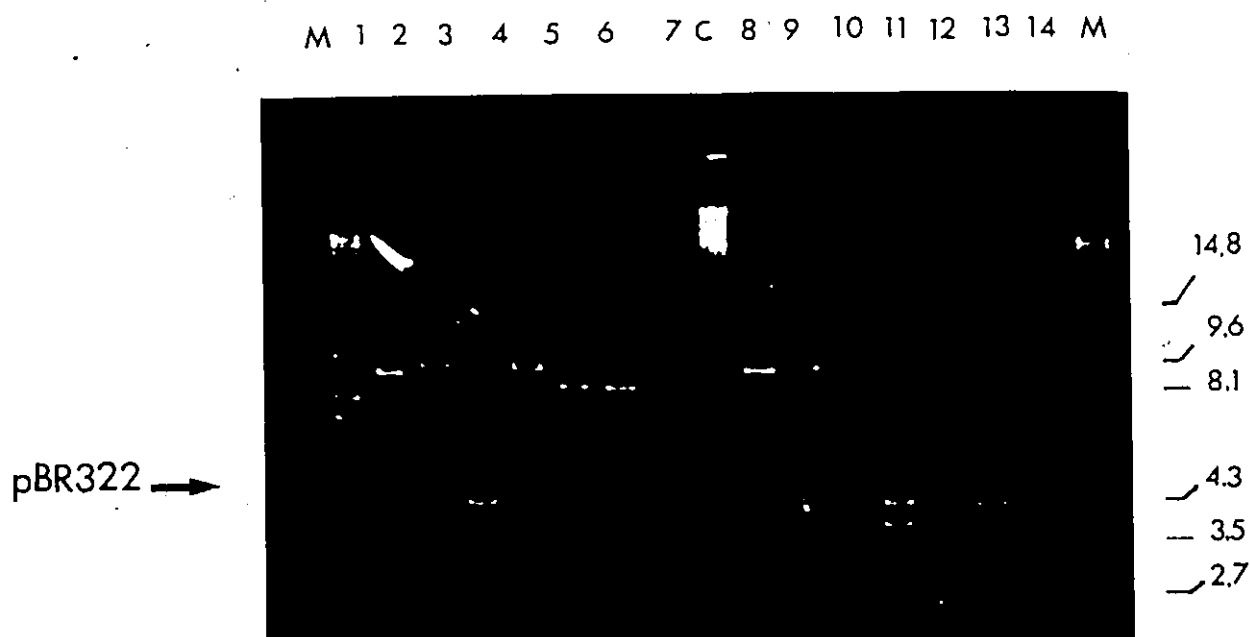


Figure 3.1 The set of soybean Pvu II recombinant plasmids digested with Pvu II to recover the cloned fragments; 9.6 Kb (lanes 2,4,9), 9.4 Kb (lane 8), 9.0 Kb (lane 1), 8.6 Kb (lanes 5, 6), 4.1 Kb (lane 10), and 3.5 Kb (lane 12). Lanes M and M contain EcoR I, Hind III digest of lambda, lane C contains Pvu II digest of soybean cpDNA. Lanes 11 and 13 are probably nuclear contaminants. The size range in Kb is at right.

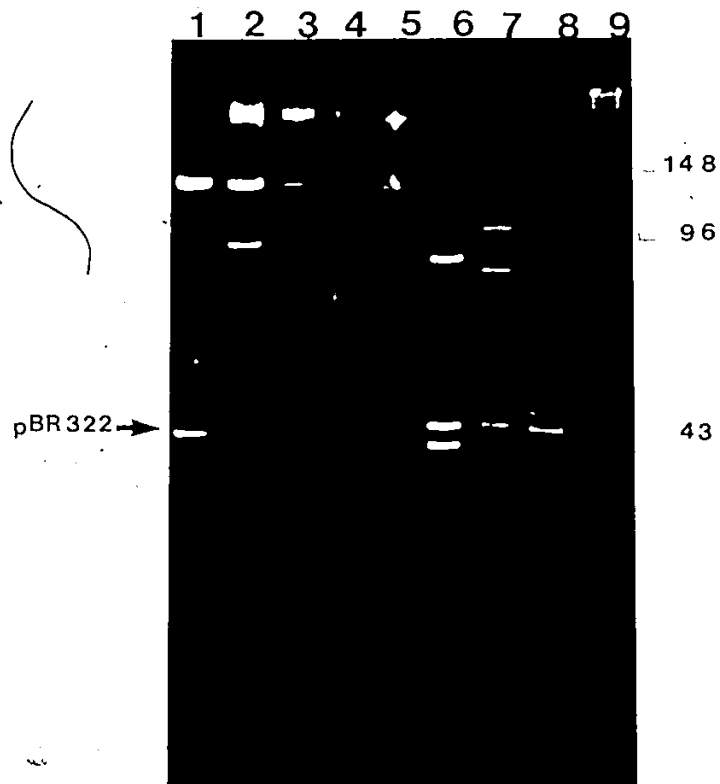


Figure 3.2 Analysis of the selected recombinant clones tentatively identified by the electrophoretic mobilities of the supercoiled plasmid, carry 14.8 Kb Pvu II fragment. Lane 1 (14.8 Kb), lane 2 (14.8, 9.4 Kb), lane 3 (14.8, 9.4 Kb), lane 4 (14.8, 8.6 Kb). Lanes 4, 6, 7, and 8 carry background. Lane 9 is collection of Pvu II digest soybean cpDNA.

Table 3.1 Cloning of soybean cpDNA fragments;
Set of Pvu II inserts in
pBR322 (4.36 Kb.)

Fragments obtained upon digestion	Cloned fragments	Clone designation
41.0		
34.0		
14.8	+	pBP3
9.6	+	pBP4
9.4	+	pBP5
9.0	+	pBP6
8.6	+	pBP7
8.1		
4.1	+	pBP9
3.5	+	pBP10
2.7 (2X) A, A'		

Sizes of the fragments are in Kb.

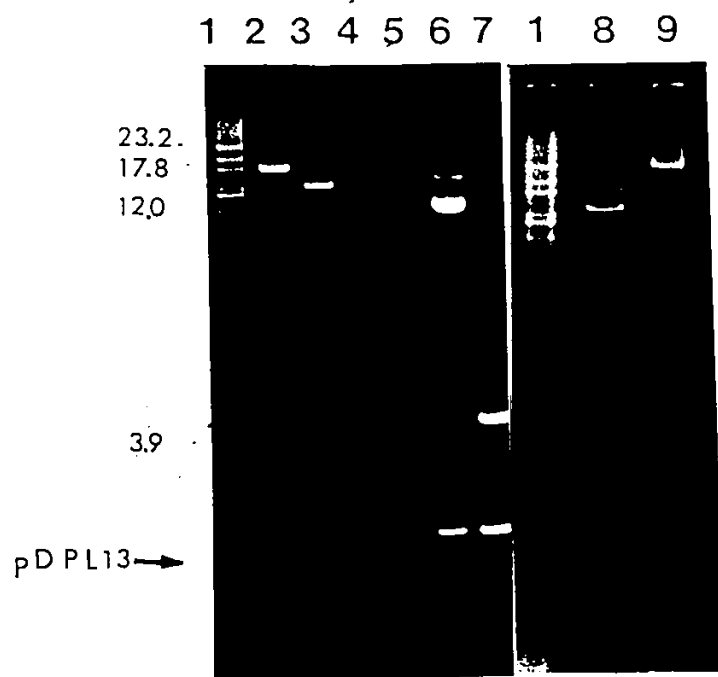


Figure 3.3 The soybean cpDNA - Sac I clones, digested with Sac I to recover the inserted fragments. Lane 1 - Sac I digestion product of cpDNA, lane 2 (17.8 Kb), lane 3 (14.5 Kb), lane 4 (13.6A), lane 5 (13.6A'), lane 6 (12.0), lane 7 (3.9 Kb), lane 8 (probably nuclear contaminant). Size estimation scale is at left.

Table 3.2 Cloning of Soybean cpDNA fragments;
Set of Sac I inserts in
pDPL13 (2.5 Kb.)

Fragments obtained upon digestion	Cloned fragments	Clone designation
23.2	+	pDS1
19.8		
17.8	+	pDS3
14.5	+	pDS4
13.6 A	+	pDS5
13.6 B	+	pDS6
12.0	+	pDS7
5.6		
4.2		
3.9 (2X) A, A'	+	pDS10
3.8	+	pDS11
3.5 (2X) A, A'		
2.8		
1.8 (2X) A, A'		
1.8 B		
1.3		

Sizes of the fragments are in Kb.



Figure 3.4 Preparative gel electrophoresis of Sac I digestion products of soybean cpDNA on 0.7% agarose. Some of the Sac I fragments are labeled.

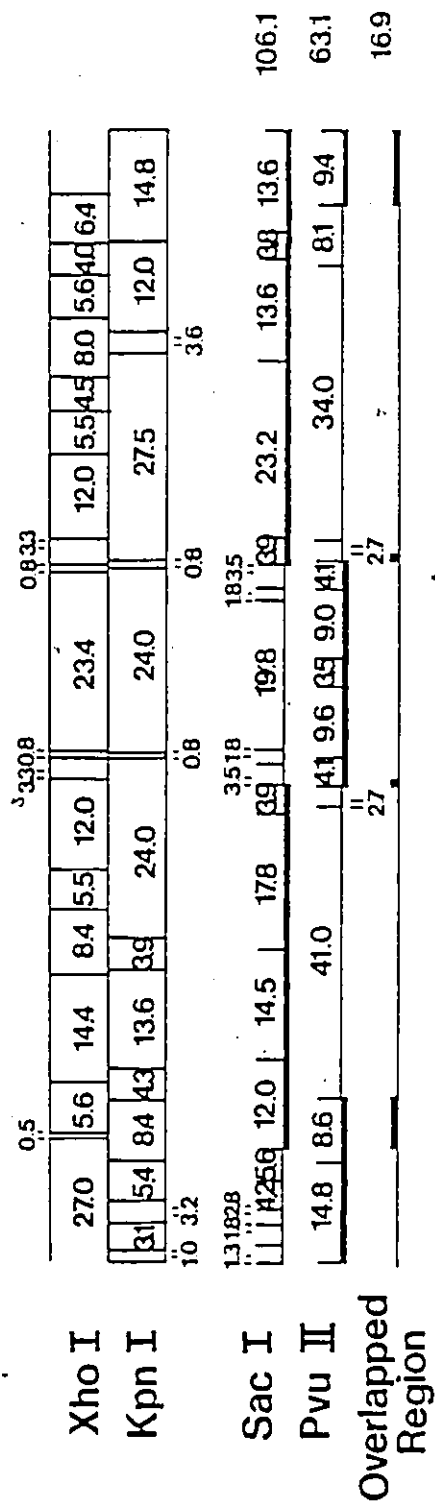


Figure 3.5 Schematic representation of the soybean clone bank. The Pvu II and Sac I cloned fragments are underlined. The soybean cpDNA regions represented in both Pvu II and Sac I clones are represented by closed bars. The total length (in Kb) of soybean cpDNA present in Pvu II and/or Sac I, is at right.

3.5 DISCUSSION

The set of Pvu II and Sac I cpDNA fragments were selected because their restriction sites were mapped and appropriate vectors containing these insertion sites were available. pBR322 (Bolivar et al, 1977) is a well-established cloning vector. Its Pvu II site is located approximately 470bp from the origin of replication. pDPL13 (2.5 Kb) a deletion derivative of pBR322, lacks the tetracycline (Tet) gene but carries a cleavage site of 13 restriction enzymes in a 79 bp polylinker (Pulleyblank - personal communication). In both cloning vectors, the insertion sites are outside the drug markers, hence insertional inactivation technique of cloning can not be achieved. This, however, did not restrict the construction of a clone library because of the high insertion and transformation frequencies observed.

The construction of a clone bank of nuclear or organellar genomes is an important step in studying the gene structure and expression. A Clone bank allows the isolation of a particular fragment of a specific gene otherwise available in small proportions and permits the analysis of gene expression in vitro or in special host cells.

To date, (Coen et al, 1977, Bedbrook et al, 1979, Link and Bogorad 1980, Gatenby et al, 1981, Steigler et al, 1981) more reports have accumulated on cloning and studying a particular region of cpDNAs than the construction of the

entire genome clone banks. Only few reports - to our knowledge - of comprehensive clone bank have been published so far: Chlamydomonas reinhardtii (Rochaix, 1977), mung bean, pea, spinach (Palmer and Thompson, 1981, tobacco (Zhu et al, 1982), broad bean (Ko et al, 1983). All the above libraries, except broad bean and mung bean are deficient in some portion of the chloroplast genome.

The present study is the first clone library of a major legume crop, soybean. Results show that the complete cpDNA is represented in 15 recombinant clones of pBR322 and pDPL13. Eight Sac I fragments ranging from 23.2 Kb to 3.8 Kb were inserted in pDPL13 and a set of seven Pvu II fragments ranging from 14.8 to 3.5 Kb were cloned in pBR322.

Approximately 17.0 Kb of the chloroplast genome is shared between the two sets of clones. The frequency of cloning beyond 15 Kb decreased with increasing size. Recombinant DNA molecules measuring more than 30 Kb are difficult to clone probably due to the inability to introduce them into bacterial cells by presently known methods.

Two of the selected Pvu II fragments, 8.1 Kb and 2.7 Kb could not be cloned despite all attempts. The corresponding regions, however, were cloned successfully by introducing Sac I fragments in pDPL13 (Fig. 3.3 and Table 3.2). The 2.7 Kb Pvu II is contained in the chimeric plasmid pDS10, and 8.1 Kb Pvu II is shared between Sac I recombinant clones,

pDS5, pDS6 and pDS11. Palmer and Thompson (1981a) could not clone 12.2 Kb Pst I fragment of pea chloroplast DNA and ascribed the failure to the sequence of the insert which they believed might interfere with the cloning process or else might be lethal to the host.

The inability to recover recombinant clones containing the moderate sized Pvu II inserts (8.1 Kb or 2.7 Kb) suggests that size of the foreign DNA may not be the only limiting factor for efficient cloning. But these Pvu II regions were stably cloned as part of Sac I fragments. This implies that the sequences of the inserts, at least in our case, are not inhibitory to either the cloning process or bacterial machinery. However, this does not preclude the possibility that the altered transcription frame of the recombinant plasmid may either result in a transcription or protein product, deleterious for the proper maintenance of the bacterial or plasmid replication. Moreover, Remault et al (1981) found that a constitutively high transcription may interfere with plasmid DNA replication and may lead to plasmid instability. Thus the presence of open reading frames or promoters near the Pvu II insertion site may influence the transcription rate of genes(s) product(s) transcribed by the foreign DNA fragment introduced at Pvu II site. Although no information is available regarding the encoding capacity of 8.1 Kb Pvu II, it is certain that 2.7 Kb Pvu II contains at least 90% of 16S rRNA gene (chapter

3). Sequence analysis of pBR322 replication region by Sutcliffe (1979) revealed the presence of two open reading frames upstream from Pvu II site. Recently, Cesareni et al (1982) have attributed a regulating function to the product. They identified the presence of an initiation signal, 6 bp downstream from the ribosome binding site, GGAGG at 1916 bp on pMB1 DNA (pBR322 contain pMD1 type replicator, Bolivar et al 1977). This signal could elevate the transcriptional rate of the gene(s) downstream which in turn might affect plasmid survival. Interestingly, Cesareni et al (1982) have found that this gene product acts as an inhibitor of DNA replication by limiting the amount of primer available for DNA synthesis. Further, they also demonstrated that this newly identified, rop gene product acts as an inhibitor of DNA replication by limiting the amount of primer available for DNA synthesis. Twigg and Sheratt (1980) found that deletion of this region depicts an increase in copy number of the plasmid. These new findings regarding the molecular anatomy of the region near Pvu II site suggest that the presence of open reading frame signals and ribosome binding site might elevate the transcription level of the gene(s) situated downstream so as to affect the plasmid stability (Remaut et al, 1981). Alternatively, transcription or protein product formed by recombinant plasmid may affect the plasmid survival. Furthermore, the inactivation of the rop gene by insertion at Pvu II would lead to increased copy

number. The region near the vicinity of Pvu II is missing in pDPL13 (Straus, Pulleyblank - personal communication) hence the recombinant clones resulting from Pvu II cloning in pBR322 or by insertion in either permissible sites in Sac I will result in elevated copy number.

In conclusion, size of the insert fragment does affect the transformation frequency. Other possible factors which might influence efficient and stable cloning are the nucleotide sequence of the recombinant plasmid and the transcription rate of the insert (Remaut et al, 1981).

Chapter IV

STRUCTURAL ORGANIZATION OF THE rRNA CISTRON

4.1 INTRODUCTION

Loening and Ingle (1967) showed that 16S and 23S were the main high molecular weight ribosomal RNA species of the chloroplast as opposed to 18S and 25S rRNAs of eukaryotic 80S ribosomes. Chloroplast ribosomes also contain low molecular weight 4.5S and 5S RNAs (Dyer et al, 1968; Dyer et al, 1976). The 4.5S, 5S, and 23S RNAs are components of 50 S subunit of chloroplast ribosome and the 16S rRNA is a part of 30 S subunit.

Thomas and Tewari (1974) using quantitative RNA-DNA hybridization showed that cpDNA in corn, spinach, pea, Chlamydomonas and E. gracilis contained two "gene equivalents" for the chloroplast rRNAs. Pea cpDNA has been shown to contain a single set of rRNA genes instead of two by other more reliable methods. Overestimation of quantitative hybridization data may lead to incorrect interpretation (Chu et al, 1981).

The presence and detailed organization of rRNA genes was initially reported in corn by Bedbrook et al (1977). Since then the presence of ribosomal RNA genes has been shown in all the chloroplast DNAs investigated so far.

Generally two methods used to obtain these results are: Nitrocellulose filter hybridization and R-loop analysis by electron microscope. Nitrocellulose filter hybridization - also referred to as Southern hybridization, (Southern, 1975) involves the binding of probe molecules to denatured cpDNA sequences fixed on nitrocellulose paper. R-loop analysis is an elegant way to demonstrate the existence of RNA-transcribing regions, and involves the hybridization of RNA and DNA molecules under reaction conditions where RNA-DNA hybrids are more stable than DNA-DNA hybrids. Renaturation in the presence of rRNA resulted in the displacement of DNA strand where rRNA was bound (R-loops). The visualization of R-loops under electron microscope confirmed the presence of rDNA bearing region (Bedbrook et al 1977, Chu et al 1981, Koller and Delius 1980). Successful application of the above-mentioned techniques to different plant chloroplast genomes has established the presence of a single set of rRNA cistron per unit of the inverted repeat. Therefore, the cpDNAs possessing the inverted repeat (Atriplex, Palmer 1982; corn, Bedbrook et al, 1977; Chlamydomonas, Rochaix and Malnoe, 1978; tobacco, Fluhr and Edelman, 1981a) contain two sets of rRNAs as opposed to pea (Chu et al, 1981, Palmer and Thompson 1981a) and broad bean (Koller and Delius, 1980, Ko et al, 1983). Such chloroplast DNAs lack an entire segment of the repeated region and therefore show the presence of a single set of rRNA genes. The arrangement of E. gracilis

rRNA is an exception to the above-mentioned modes of organization. In the two most frequently studied strains, bacillaris and Z, there are three copies of tandemly repeated 5.6 Kb sets, each containing a 16S and 23S rRNA gene and separated by a spacer (Gray and Hallick, 1978) 1978, Rawson et al, 1978, Helling et al 1979). Few differences in the length and nucleotide sequence of the intervening spacer between the two strains have been reported (Jenni and Stutz, 1978, Wurtz and Breton, 1981). More variation in the structural organization and redundancy of chloroplast rDNA cistrons occurs within the Euglena gracilis spp. than among the higher plant cp rDNAs examined. In all higher plant cpDNAs examined, rDNAs are part of the inverted repeat structure.

Detailed restriction analysis of the rDNA region has assisted in exploring the structural organization and the sequential order of the ribosomal RNAs. From the studies of various workers (Bedbrook et al, 1977, Whitfield et al, 1978, Malnoe and Rochaix, 1978, Gray and Hallick 1979, Delius and Koller, 1980) it is confirmed that the order of rRNA genes is 5' - 16S-spacer-23S spacer-5S - 3', similar to that found in E. coli (Lund et al, 1976) and blue-green algae, Anacystis nidulans (Tomika et al, 1981). Bedbrook et al (1977) determined the sizes of 16S and 23S rDNAs by measuring the lengths of RNA-DNA hybrids from electron micrographs. These estimations are lower than the actual

size determination of 16S rDNA and 23S rDNA by sequence analysis. (Schwartz and Kossel, 1980; Edwards and Kossel, 1981). The E. coli 16S rRNA has at least 72% homology with E. gracilis, N. tabacum, and Z. mays, whereas N. tabacum and Z. mays have 96% homology. The chloroplast 23S rRNA genes from corn and tobacco show 92% sequence homology between themselves, and approximately 67% with the E. coli 23S rRNA.

Downstream from 23S rDNA there is a brief spacer followed by a 4.5S coding region (Edwards et al 1981, Takaiwa and Sugiura 1980). This region has at least 67% homology with 3' end of E. coli 23S rDNA (Machatt et al 1981) but the 23S - 4.5S spacer sequence has no equivalent in the bacterial gene. The formation of 4.5S rDNA may have been the result of an insertion at the 3' end of 23S RNA gene (Edwards et al, 1981). A defined role has not been attributed to the 4.5S rRNA found in the 50S ribosomal subunit (Whitfield et al, 1978). The 3' end of 4.5S rDNA shows extensive homology with 5' end of 23S rDNA (Takaiwa and Sugiura, 1980, Edwards et al, 1981) In association with 23S rRNA it may form cruciform structures analogous to that formed by E. coli precursor rRNA prior to processing, and thus could play a role in chloroplast rRNA processing.

The 5S rRNA gene is located downstream from the 4.5S rRNA by 256 bp in tobacco (Takaiwa and Sugiura 1980). They have found the "Pribnow box" and the "-35 region" located in the 4.5 - 5S rRNA spacer which could function as a signal

for the initiation of transcription of 5S rRNA gene. This study suggests that 5S rRNA may be transcribed separately from the 16S - 23S - 4.5S' cluster.

The 16S - 23S rRNA spacer has attracted a lot of attention due to its structure. This region, completely sequenced in corn (Koch et al, 1981), tobacco (Takaiwa and Sugiura, 1982b), and Euglena (Graf et al, 1980, Orozco et al, 1982), reveals the presence of two tRNA genes, tRNA^{Ile} and tRNA^{Ala}. Corn and tobacco 16S - 23S spacer is much larger than the Euglena spacer due to the presence of intron in the tRNA genes (Koch et al, 1981 and Takaiwa and Sugiura, 1982b).

Sequence analysis and other studies (Hartley, 1979) indicate that spacer DNA is transcribed as the part of rRNA operon transcriptional unit which is subsequently processed to generate rRNA and tRNA species

Chloroplast ribosomal RNA genes, as indicated above, are among the more thoroughly investigated regions of the chloroplast chromosome. To compare these findings with the soybean, a detailed study of the cp rRNA genes was undertaken. Localization, relative orientation, and copy number of the ribosomal RNA genes was determined by Southern hybridization. Portions containing the major rRNA genes (16S and 23S) were cloned in suitable vectors and used for fine mapping. Information regarding the direction of transcription was collected from DNA hybridization with

filter fixed cp rRNAs. Chloroplast rRNA regions of soybean and tobacco (Takaiwa and Sugiura, 1982) have been compared in overall structure and nucleotide sequence homologies.

4.2 MATERIAL AND METHODS

4.2.1 Extraction of chloroplast and plasmid DNA

Chloroplast and plasmid DNAs were obtained as described earlier (chapter 1 and 2).

4.2.2 Isolation of soybean chloroplast rRNA

Chloroplast rRNA was prepared either from green cotyledons or leaves. Eight day old cotyledons or 7 - 10 day old leaves were found satisfactory for all the different species of rRNA. All subsequent steps were carried out at 4°C unless otherwise mentioned. The harvested material was homogenized in prechilled buffer containing 0.5 M sucrose, 100 mM Tris - HCl, pH 8.0 25 mM MgCl₂ and 1 mM spermidine, in Waring blender for 5 sec at low speed. The homogenate was filtered through 4 layers of cheesecloth, a single layer of 50 µm, and 25 µm nylon cloth. The filtrate was centrifuged at 1020 g for 5 min. The chloroplast pellet was resuspended in 100 mM Tris - HCl, 8.0, 25 mM MgCl₂, 1 mM spermidine, 4% Triton x 100 and 2% SDS for 30 min. and the lysate was centrifuged at 12,000 g for 15 min. to recover the supernatant. To this, NaCl was added to the final concentration of 0.5 M and mixed gently. The saline

supernatant was extracted twice with water-saturated phenol at room temperature and the aqueous phase was separated by centrifugation at 15,000 g. Approximately 2.5 volumes of chilled ethanol were added and the contents stored at -20°C for 4 - 5 hrs. The RNAs harvested by centrifugation at 15,000 g x 15 min.

4.2.3 Separation of various species of soybean cpRNA

High molecular weight rRNA (16S and 23S) were separated on a 5-30% sucrose gradient in 5 mM Tris, pH 8.0, 5 mM MgCl_2 , 1 mM spermidine at 40,000 rpm for 7 hrs at 4°C . Fractions containing 16S, 23S were separately collected and ethanol precipitated. The low molecular weight rRNAs were separated on 12% acrylamide-bisacrylamide gel. The 4.5S and 5S rRNAs were purified out of the gel, extracted with phenol and recovered by ethanol precipitation. All the soybean cpRNAs were kept lyophilized at -20°C .

4.2.4 Separation of E. coli 16S and 23S rRNA

E. coli 16S and 23S rRNA were purchased from Boehringer Mannheim and separated on 5 - 20% linear sucrose gradient prepared in 20 mM Tris - HCl, pH 7.4, 200 mM potassium acetate, 2 mM EDTA at 35,000 rpm in Beckman SW41 rotor for 7 hours at 4°C . Fractions containing 16S and 23S rRNA were separately pooled and the respective rRNAs were recovered by ethanol precipitation.

4.2.5 Restriction, Digestion and Fractionation

The cpDNA was digested with restriction enzymes according to the suppliers' instructions and the cleaved products were fractionated on normal or low melting agarose as described earlier (page 15).

4.2.6 Recovery of labeled DNA fragments from gel

After the fragments were fractionated on agarose containing ethidium bromide, the gel was visualized under UV illumination. Labeled fragments were identified by hand-held geiger counter and the bands were cut, transferred to Eppendorf tubes, diluted to obtain 0.2% agarose concentration and denatured at 100 C for 15 min.

4.2.7 Southern and Northern Blots

DNA fragments over 1 Kb were transferred to nitrocellulose paper according to a procedure modified after Southern (1975). Collection of fragments below 1 Kb were transferred to nitrocellulose paper by the procedure generally referred as "Northern" (Thomas, 1980). The gel fractionated rRNAs were transferred to nitrocellulose paper by both Southern and Northern procedures.

4.2.8 Radioactive labeling

rRNA labeling: About 10 ug of RNA in total volume 40 ul was alkali hydrolyzed for four min. and neutralized with prechilled 260 ul of 100 mM sodium acetate and 100 mM acetic acid. Hydrolyzed rRNA fragments (80 - 100 bp) were ethanol precipitated and resuspended in bidistilled water to yield 1 ug/ul concentration. These fragments (1 ug) were end labeled with 1 - 5 units of polynucleotide kinase in 10 ul reaction mixture containing 100mM Tris-HCl, 20mM DTT, 120mM MgCl₂ and 30 uCi γ -³²P ATP at a specific activity of 4000 Ci/mmmole for 45 min. at 37 °C. Reaction was terminated by the addition of 25 mM EDTA and 1% SDS and rRNA probes were precipitated with ethanol. The kinase charging generates a total activity of higher than 5×10^7 cpm.

DNA Labeling; The DNA was labeled either by filling recessed 3' end by the Klenow fragment of E. coli DNA polymerase or by nick translation (Maniatis et al, 1975). Between 0.4 ug and 2 ug of DNA was labeled by 1.0 unit of Klenow fragment in 50 mM Tris - HCl, pH7.2. 10 mM MgSO₄, 0.1 mM DTT 50 ug/ml BSA at room temperature for 30 min.

4.2.9 Hybridization

RNA-DNA and DNA-DNA hybridization were carried out in the manner described in chapter 1 (page 17).

For DNA - RNA Northern hybridization The method is a slight modification of that of Thomas (1980). The filters

were presoaked in 50% formamide, 5 x SSC, (Saline sodium citrate) citrate; 1X SSC is 150 mM NaCl, 15mM sodium citrate). denatured calf thymus DNA at 100 ug/ml and 5 X Denhardt's (1966) in heat sealable bags for 8 - 10 hours at 42°C. The hot probes were added to the same buffer after the pretreatment period. The bags containing filters were incubated for 20 - 24 hours at 42°C.

4.2.10 Washing

The hot hybridization solution was carefully transferred to falcon 15 ml disposable tubes. Washing treatment depended upon the type of hybridization reaction. DNA - DNA Southern and DNA - RNA Northern filters were washed by a similar procedure. Filters were washed with three changes of 2 x SSC, 0.2% SDS for 15 min. each at room temperature and then washed with two changes of 0.1 x SSC, 0.2% SDS at 60 - 68°C for two hours.

RNA - DNA Southern filters were washed with two changes of 2 x SSC, 0.2% SDS at room temperature, incubated with RNase (50 ug/ml) in 1 x SSC, 0.2% SDS at 37 °C and then washed in six changes of 0.1 x SSC 0.2% SDS for 30 min. each.

4.2.11 Autoradiography

The hot blots were exposed to Fuji X-Ray JX (8' x 10') film at -80°C with Dupont cronex-Hi plus screens for 6 hours - 3 weeks.

4.3 RESULTS

Ribosomal RNA genes were localized on soybean chloroplast DNA by hybridizing ^{32}P ATP labeled 16S and 23S rRNA from E.coli and tobacco to filter bound single and double digests of restriction enzymes; Kpn I, Pvu II, Sac I, Xho I. The results (Table 4.1) based upon the autoradiograph data (Fig. 4.1) indicate that rRNA genes are present within the inverted repeat. The 16S rRNA from both E. coli and tobacco hybridized with equal intensity to bimolar 2.7 Kb Pvu II, 3.9 Kb Sac I and 12.0 Kb Xho I fragments within the repeated region. The minimum size of single digest carrying the 16S rRNA gene was 2.7 Kb Pvu II. Table 4.1 shows that the radioactive 16S rRNA did not hybridize to adjacently lying 4.1 Kb, 34.0 Kb, and 41.0 Kb Pvu II fragments, indicating that over 95% of 16S rRNA gene resides within the 2.7 Kb Pvu II region. Sac I 3.9 Kb fragment carries the entire length of the 16S rRNA gene. Similarly the 23S rRNA from E. coli and tobacco hybridized to bimolar 0.8 Kb Kpn I; 4.1 Kb Pvu II; 3.5 and 1.8 Kb Sac I; 3.3 and 0.8 Kb Xho I fragments. Other fragments which hybridized were 27.5 and 24.0 Kb Kpn I and 23.4 Kb Xho I.

fragments. With the aid of the hybridization data (Fig. 4.1 and Table 4.1) and physical map (Fig. 4.2), it could be predicted that over 95% of the 23S^o rRNA gene is contained in the 4.1 Kb Pvu II fragment.

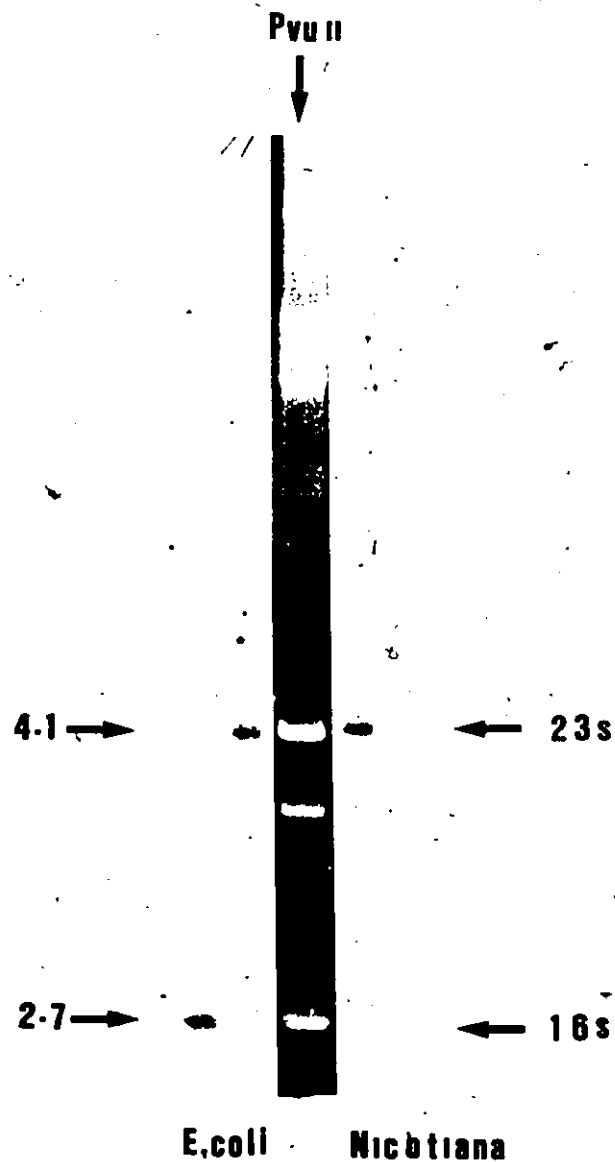


Figure 4.1 Hybridization of γ -³²P ATP ribosomal RNA probes with digestion products of chloroplast DNA with restriction endonuclease Pvu II (lane c). 16S (lane a) and 23S (lane b) ribosomal RNA from *E. coli*. Lanes d and e are 23S and 16S ribosomal RNA from *Nicotiana tabacum* respectively.

Table 4.1 Hybridization of radiolabeled ribosomal RNAs (16S and 23S) from E. coli and N. tabacum to the restriction enzyme digests of soybean chloroplast DNA.

Restriction Enzymes	γ - ³² P ATP rRNA probes			
	<u>N. tabacum</u>		<u>E. coli</u>	
	16S	23S	16S	23S
Xho I	12.0	23.4, 3.3, 0.8	12.0	23.4, 3.3, 0.8
Xho I/Sac I	3.1	2.5, 1.8, 0.8	3.1	2.5, 1.8, 0.8
Sac I	3.9	3.5, 1.8	3.9	3.5, 1.8
Xho I/Pvu II	2.0	2.6, 0.8	2.0	2.6, 0.8
Pvu II	2.7	4.1	2.7	4.1
Xho I/Kpn I	12.0	2.6, 0.7	12.0	2.6, 0.7
Kpn I	24.0	24.0, 0.8	24.0	24.0
Kpn I/Pvu II	2.7	2.0, 1.3, 0.8	2.7	2.0, 1.3
Kpn I/Sac I	3.9	1.9, 0.8	3.9	1.9, 0.8
Pvu II/Sac I	2.7	3.5	2.7	3.5

Sizes of the fragments are in Kb.

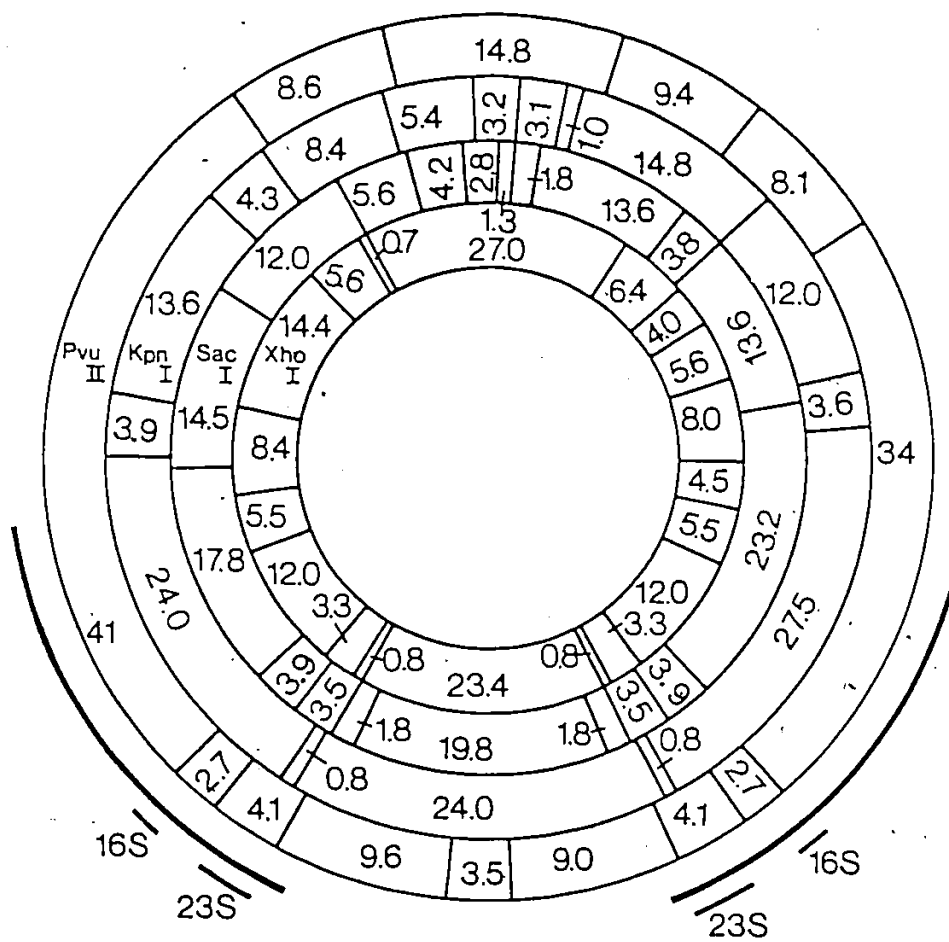


Figure 4.2 Localization of 16S and 23S rRNA genes on the restriction endonuclease cleavage site map of soybean chloroplast genome. The inverted repeat is represented by the long thick arcs.

4.3.1 Spatial arrangement of 16S and 23S rRNA genes

All the cpDNA fragments hybridizing to the 16S and 23S rRNA genes were present at the ends of the inverted repeat region. The spatial arrangement of 4.1 Kb Pvu II and 2.7 Kb Pvu II fragments (Fig. 4.2) dictated the sequence of rRNA genes as 16S - 23S within the repeat unit. The size estimation of 16S and 23S by Bedbrook et al (1977) and Chu et al (1981) allow only one copy of 16S and 23S to be present in the respective fragments. As 16S and 23S rRNA probes hybridized to bimolar fragments encompassed within the inverted repeat, two copies of rRNA cistrons must be present per chloroplast chromosome.

4.3.2 Cloning

Pvu II fragments 2.7 Kb and 4.1 Kb were selected for cloning in Pvu II site of pBR 322. As mentioned in chapter 2, the 4.1 Kb Pvu II was cloned in pBR322 but the 2.7 Kb Pvu II could not be cloned in the chosen vector in spite of several attempts. Hybridization data (Table 4.1) suggested that the 3.9 Kb Sac I fragment contained the 16S rDNA in toto, hence this fragment was cloned in pDPL13. The recombinant plasmids pDS10 and pBP9 contained 16S and 23S rRNA genes respectively and were therefore utilized for fine mapping of rDNA cistron.

4.3.3 Fine mapping

The recombinant plasmids were screened with over 20 different (tetra, penta or hexanucleotides) restriction endonucleases in order to find enzymes which generate a simple, easily discernable restriction pattern. Twelve enzymes were chosen to fine map the rDNA region.

4.3.4 Mapping strategy

The mapping scheme is different than the one described in Chapter 1. The information regarding the discrete order of restriction sites was obtained by a rather inexpensive method of single, double and reciprocal digestions, which is illustrated by positioning restriction sites of endonucleases *Av* II and *Sma* I on pDS10.

The information regarding the number and relative positioning of *Kpn* I, *Pvu* II and *Xho* I restriction sites was based upon the cleavage map of soybean chloroplast DNA (Chapter 1). *Sma* I restriction sites were aligned by double and reciprocal digestions. *Sma* I digestion yields three fragments; 2.4 Kb, 0.96 Kb and 0.68 Kb (Table 4.2). Comparison of *Sma* I - *Xho* I double digests with single digests of *Sma* I and *Xho* I, imply that the fragment (0.96 Kb) generates 0.86 Kb plus 0.1 Kb subfragments upon *Sma* I - *Xho* I double digestion, and 0.78 Kb *Xho* I results in 0.68 Kb and 0.1 Kb subfragments upon double digestion indicating that 0.7 Kb *Xho* I overlaps 0.96 Kb and 0.68 Kb *Sma* I. The

largest 2.4 Kb Sma I fragment was retained during the Sma I - Xho I double digestion, hence establishing that 2.4 Kb Sma I is completely contained within the 3.1 Kb Xho I fragment (Fig. 4.3).

Table 4.2 Fine mapping of the 3.9 Kb Sac I insert;
Positioning of Ava II, Sma I sites.
Molecular sizes (Kb) of sub-fragments of
the 3.9 Kb Sac I insert produced by
single and double digestion of Ava II,
Sma I and Xho I.

Xho I	Sma I/ Xho I	Sma I	Ava II/ Sma I	Ava II	Ava II/ Xho I
3.1	2.40	2.40	1.35	1.80	1.80
0.78	0.86	0.96	1.00	1.35	1.35
	0.68	0.68	0.80	0.90	0.78
	0.10		0.68		
			0.20		
3.88	4.04	4.04	4.05	4.05	3.93

The sizes of the fragments were determined by comparing
their electrophoretic mobilities with digests on
1.5-2.0 % agarose.

Sizes of the fragments are in Kb.

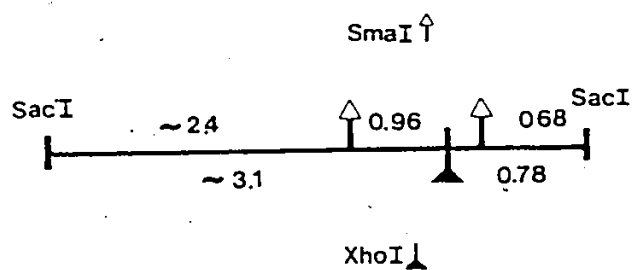


Figure 4.3 Positioning of the Sma I and Xho I cleavage sites on the 3.9 Kb Sac I insert.

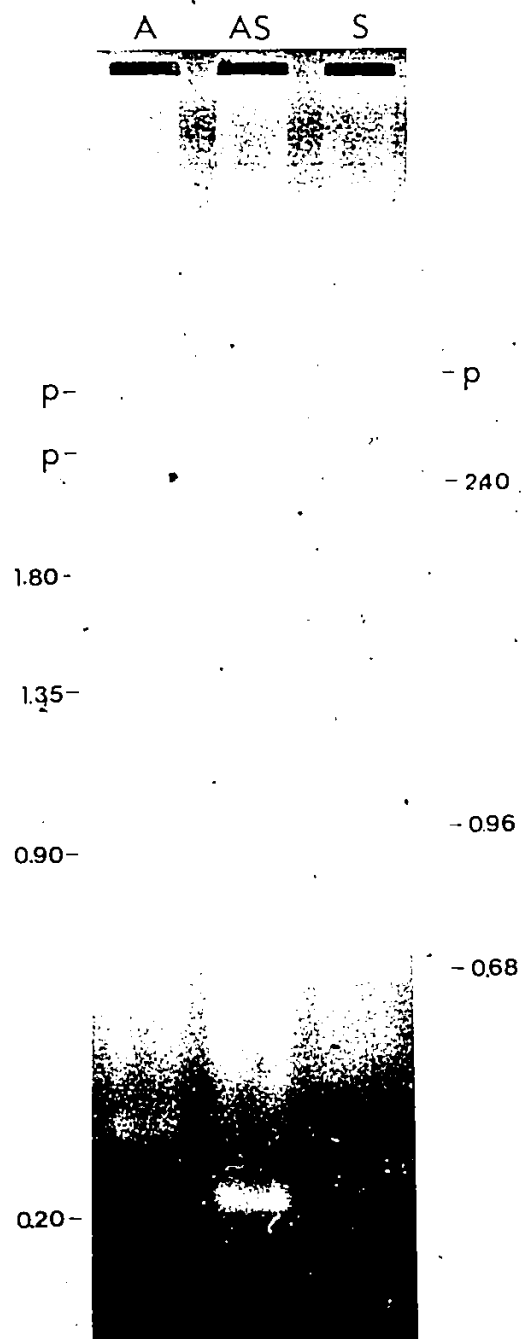


Figure 4.4 Digestion pattern of 3.9 Kb Sac I insert with Ava II (A), Ava II and Sma I (AS) and Sma I (S) on 1.6% agarose gel. The sizes are quoted in Kb. The fragments marked "p" are the partial digests.

Table 4.3 Fine mapping of the 3.9 Kb Sac I insert;
Positioning of Ava II Sma I sites by
reciprocal digestion.

Ava II fragments	Cleaved with Sma I	Sma I fragments	Cleaved with Ava II
1.80	1.00, 0.80	2.40	1.35, 1.00
1.35	1.35	0.96	0.80, 0.20
0.90	0.68, 0.20	0.68	0.68

Sizes of the fragments are in Kb.

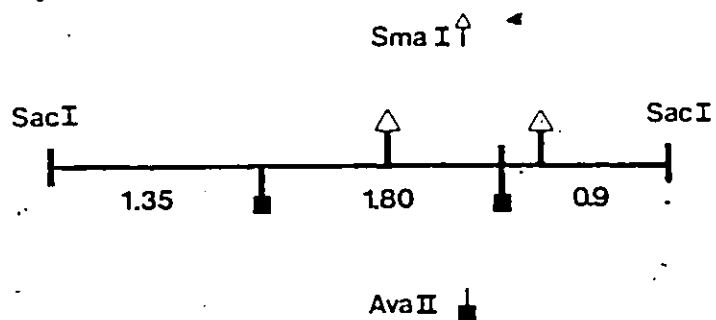


Figure 4.5 Mapping of cleavage sites of restriction endonuclease Ava II on the Sma I subfragments of the 3.9 Kb Sac I insert.

Ava II restriction sites were mapped on the Sma I fragments by double and reciprocal digestions Fig. 4.4 shows that Ava II digestion of the 3.9 Kb Sac I resulted in three fragments; 1.80 Kb 1.35 Kb and 0.9 Kb, Of these, 1.35 Kb Ava II was retained in Ava II - Sma I double digestion suggesting that 1.35 Kb Ava II is contained in the 2.4 Kb Sma I fragment. Similarly, 0.68 Kb Sma I was left uncut in Ava II - Sma I digestion indicating that it might be carried within 1.8 Kb or 0.90 Kb Ava II fragments. This uncertainty was resolved by reciprocal digestion. Ava II and Sma I single digest were recovered from the gel by electroelution and extracted twice with phenol to remove agarose. Ava II fragments were digested with Sma I and viceversa. The results of Ava II fragments redigested with Sma I indicates that the 1.35 Kb Ava II fragment remains intact; 1.80 Kb Ava II yields 1.00 Kb and 0.80 Kb subfragments, and the 0.9 Kb Ava II generates 0.68 Kb and 0.20 Kb subfragments (Table 4.3). Likewise, analogous digestion of Sma I fragments with Ava II shows that 2.4 Kb Sac I gave 1.35 Kb and 1.00 Kb subfragments and the 0.96 Kb Sma I resulted in 0.80 Kb and 0.21 Kb smaller fragments, while the 0.68 Kb Sma I remained intact. Based upon this information, Ava II sites can be positioned on the 3.9 Kb Sac I insert in pDS10 (Fig. 4.5). Molecular weights of smaller cpDNA fragments (4.0 Kb) were recalculated on higher percentage of agarose (1.5% - 2.0%). 3.9 Kb Sac I fragment was recalibrated by adding molecular

weights of the subfragments. The corrected molecular size is now 3.95 Kb.

Using the above-mentioned principle, the cleavage sites of other restriction enzymes were positioned (Tables 4.4, 4.5 and Fig. 4.6, 4.7). No Hind III and Kpn I sites were found on the Sac I insert (pDS10) and no sites for Pst I and Sal I were present on either of the inserts (Fig. 4.8).

Table 4.4 Fine mapping of 3.9 Kb Sac I insert;
Digestion products of cpDNA 3.9 Kb
Sac I insert with different
restriction enzymes (single and double
digests).

Restriction Enzymes	Digestion products	Total
Xho I	3.10, 0.78	3.88
Xho I/BamH I	2.90, 0.78, 0.12	3.80
BamH I	2.90, 0.90	3.80
Xho I/Bgl II	2.60, 0.78, 0.40	3.78
Bgl II	3.40, 0.40	3.80
EcoR I	1.40, 1.02, 0.80, 0.77	3.99
Xho I/Hind III	3.10, 0.78	3.88
Hind III	NP	
Xho I/Kpn I	3.10, 0.78	3.88
Kpn I	NP	
Xho I/Pvu II	2.00, 1.22, 0.60, 0.15	3.97
Pvu II	2.60, 1.22, 0.15	3.97
Xho I/Pst I	3.10, 0.78	3.88
Pst I	NP	
Xho I/Sal I	3.10, 0.78	3.88
Sal I	NP	

Sizes of the fragments are in Kb.
NP - no digestion product obtained.

Table 4.5 Fine mapping of 4.1 Kb Pvu II insert.
Digestion products of cpDNA Pvu II
insert with different restriction
enzymes (single and double digests).

Restriction Enzymes	Digestion Products	Total
Xho I	2.65, 0.78, 0.68	4.11
Xho I BamH I	2.65, 0.78, 0.34 (2X)	4.11
BamH I	3.42, 0.34 (2X)	4.10
Xho I/Bgl II	1.80, 1.30, 0.75, 0.30	4.15
Bgl II	3.80, 0.30	4.10
Xho I/EcoR I	2.65, 0.75, 0.60, 0.10	4.10
EcoR I	3.30, 0.71, 0.10 ^a	4.11
	0.29, 0.22	4.09
Hind III	1.80, 1.10, 0.89, 0.29	4.08
BamH I/Kpn I	2.10, 0.75, 0.60	4.13
	0.34 (2X)	
Kpn I	2.10, 1.30, 0.75	4.15
Sac I	3.45, 0.51, 0.15 ^b	4.11
Xho I/Sma I	2.64, 0.68, 0.51, 0.26	4.09
Sma I	2.90, 1.20	4.11
Xho I/Sal I	2.65, 0.78, 0.68	4.11
Sal I	NP	
Xho I/Pst I	2.65, 0.78, 0.68	4.11
Pst I	NP	

Sizes of the fragments are in Kb.

a - theoretically deduced.

b - checked on 2.0% agarose.

NP - no digestion product obtained.

Table 4.6 Localization of 16S rRNA gene;
Hybridization of γ - ^{32}P labeled 16S rRNA
from E. coli and Glycine max to
sub-fragments of Sac I insert.

Restriction digest	16S rRNA
	<u>E. coli/G max</u>
Ava II	1.85, 1.35
BamH I	2.90
EcoR I	1.40, 0.77
Pvu II	2.60, 1.22
Sma I	2.40, 0.96
Xho I	3.10

Sizes of fragments are in Kb.

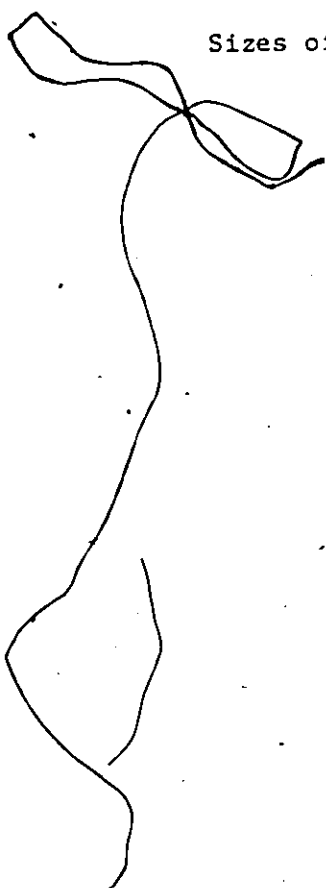


Table 4.7 Localization of 23S rRNA gene;
Hybridization of γ - ^{32}P labeled
23S rRNA from E. coli and Glycine
max to sub-fragments of Pvu II
insert (pBP9)

Restriction digest	p- 23S rRNA <u>E. coli/G. max</u>
BamH I	3.42, 0.34
EcoR I	3.30, 0.70
Hind/III	1.80, 1.10, 0.89, 0.30
Kpn I	2.10, 1.30, 0.75
Sac I	3.45, 0.51
Sma I	2.90, 1.20
Xho I	2.65, 0.78, 0.68

Sizes of fragments are in Kb.

4.3.5 Organization of rRNA genes

16S and 23S rRNAs from E. coli and soybean were individually end labeled with 32 P-ATP and T4 polynucleotide kinase according to procedure modified after Maizels (1976). The DNA fragments from pDS10 digested with Ava II, BamH I, EcoR I, Pvu II, Sma I and Xho I were fractionated by electrophoresis on agarose gels denatured and transferred directly to nitrocellulose filters. In the same manner, cleaved products of pBP9 insert with restriction enzymes EcoR I, Hind III, Kpn I, Sac I and Xho I were resolved and transferred to nitrocellulose filters. Individual 32 P labeled 16S and 23S rRNA from E. coli, soybean and tobacco were hybridized with nitrocellulose-bound DNA fragments from pDS10 and pBP9 inserts respectively. The hybridization results are compiled in Table 4.6 and 4.7 16S rRNA from both E. coli and soybean hybridized to 0.77 Kb and 1.4 Kb EcoR I, 1.22 Kb and 2.6 Kb Pvu II, 2.9 Kb BamH I; 23S rRNA hybridized to all the single digestion products of the selected enzymes except 0.1 Kb EcoR I and 0.15 Kb Sac I. The results of these hybridizations give the approximate location of 16S and 23S within the Sac I and Pvu II inserts respectively (Figs. 4.6, 4.7). Tables 4.6, 4.7 and Figs. 4.6, 4.7 depict the DNA fragments which hybridize to labeled 16S rRNA and 23S rRNA probes. 16S rRNA hybridized to 0.2 Kb EcoR I - Pvu II; 0.77 Kb, 1.4 Kb EcoR I and 2.6, 1.22 Kb Pvu II. However, it did not hybridize to 0.6 Kb Bgl II - EcoR

I, a double digest fragment adjacent to 1.4 Kb EcoR I. Autoradiographic data of 23S rRNA - DNA hybridization reveals that the probe hybridized to the entire digests of Hind III, Kpn I, Sma I, Xho I except 0.3 Kb Bgl II. Surprisingly none of the restriction enzymes employed cleave in the 1.5 Kb stretch between Bgl II and Hind III (Figs. 4.7, 4.8). Comparison of soybean total chloroplast rRNA with E. coli 16S and 23S showed that 16S and 23S rRNAs were approximately equal to 1.5 Kb and 3.0 Kb. Bedbrook et al (1977) using electron microscopy determined the sizes of 16S and 23S rRNA hybridization data and size estimation, we predict that 16S and 23S rRNA are separated by a 2.3 - 3.0 Kb spacer.

Localization of 16S and 23S rDNA on the circular map (Fig. 4.2) shows that the two sets of rRNA genes are separated by a large (116 Kb) and small (21 Kb) copy region. Moreover, the two sets of rRNA genes have an inverted orientation with respect to each other.

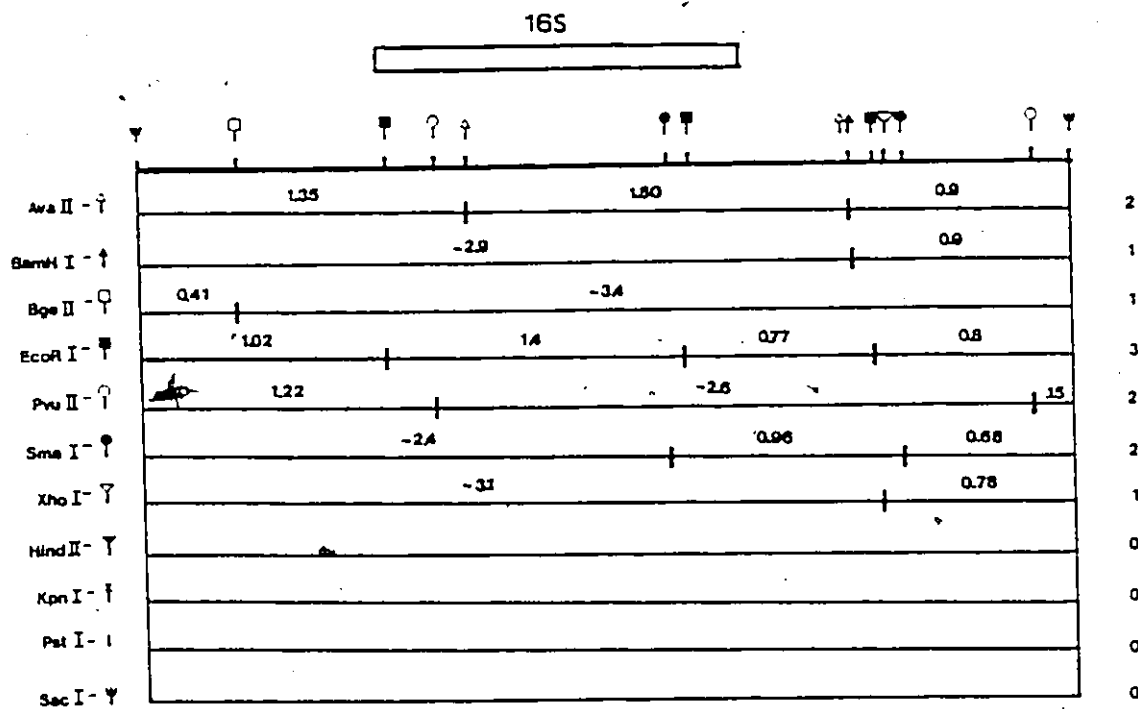


Figure 4.6 Restriction map of 3.9 Kb Sac I insert. The various endonucleases employed are shown at left. The positions of the cleavage sites of each restriction endonuclease are represented by vertical bars. The number of cleavage sites of individual endonucleases are at right. The open bar at the top shows the location of 16S rRNA gene. All the sizes are in Kb.

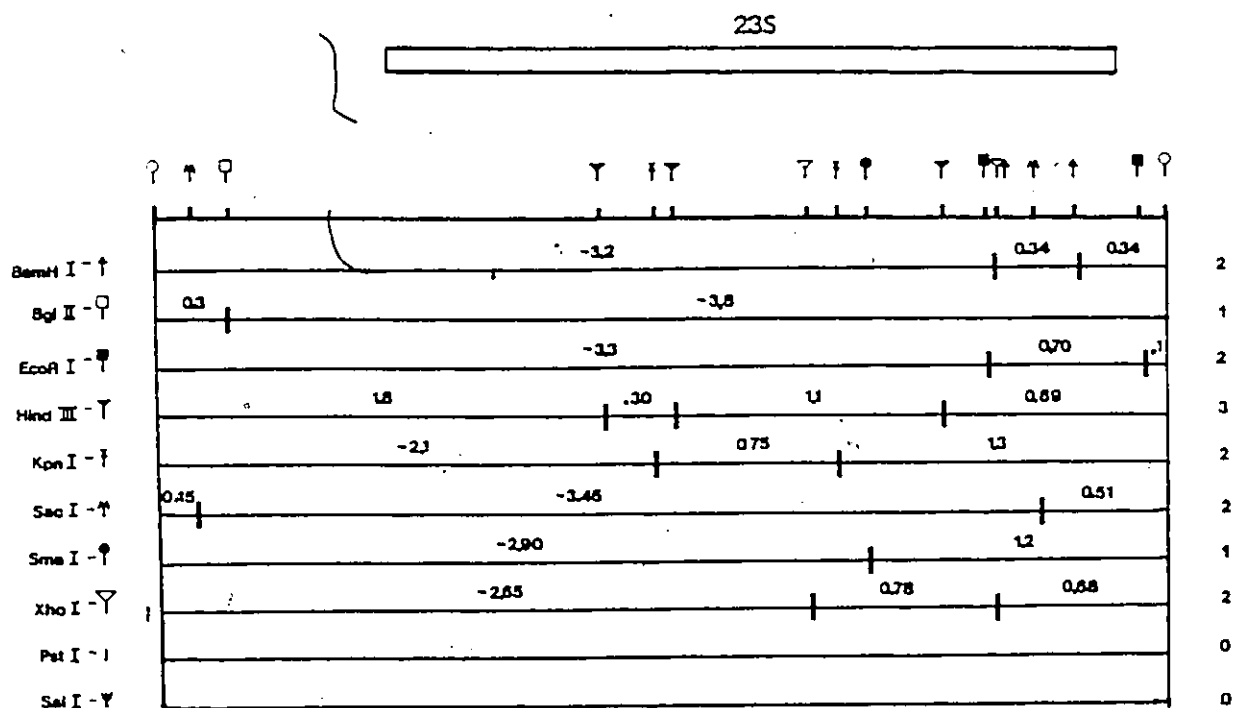


Figure 4.7 Cleavage map of 4.1 Kb Pvu II insert. Endonucleases used are shown at left. Numbers of cleavage sites of individual enzymes are given on the right. The positions of cleavage sites are represented by vertical bars. The bar at the top shows the position of 23S rRNA gene. All the sizes are in Kb.

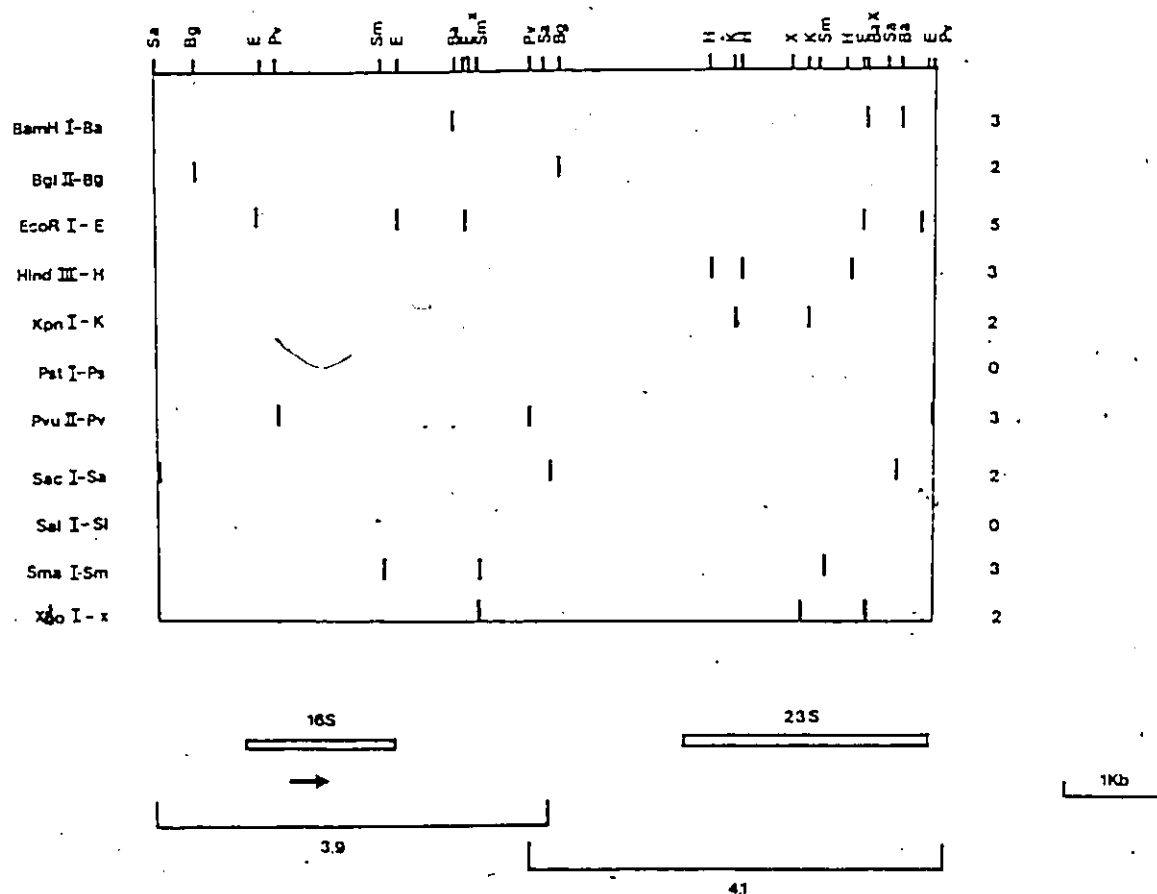


Figure 4.8: Fine mapping of the rRNA cistron by nine different restriction enzymes. The extent of the 4.1 Kb Pvu II and 3.9 Kb Sac I cloned fragments is depicted by flat brackets. The open bars represent the location of 16S and 23S rRNA gene in the cistron. Arrow shows the direction of transcription of 16S rRNA gene.

4.3.6 Homology with *E. coli* and tobacco

16S rRNA from *E. coli*, soybean and tobacco, hybridized to the same cpDNA regions with equal intensity. On the other hand, the degree of hybridization varied for 23S rRNA. *E. coli* 23S rRNA hybridized to the same cpDNA fragments as 23S rRNA from soybean and tobacco but with lowered intensity. The end labeled 5S rRNA from *E. coli* did not hybridize to filter-bound soybean chloroplast DNA fragments at either low or moderate stringency of washing conditions. Perhaps the homology between the *E. coli* 5S rRNA and soybean 5S rDNA is quite low. Similarity in hybridization pattern between 16S and 23S rRNA of *E. coli* and higher plants, soybean and tobacco indicates the apparent prokaryotic nature of the chloroplast rRNA. This suggests that *E. coli* rRNA may be employed to localize 16S and 23S rRNA regions of higher plants. The exact extent of homology between various rDNAs can be determined by sequencing studies.

4.3.7 Direction of Transcription

The direction of transcription of 16S rDNA region was established by hybridizing two fragments; Bgl II - Ava II (1.0 Kb) and Ava II - Ava II (1.85 Kb) with filter bound chloroplast rRNAs. These two fragments were eluted from the gel and extracted with phenol to remove traces of agarose. Bgl II cleaves the DNA fragment, generating 5' P tails. These extended portions were utilized as templates to add

radioactive nucleotides to 3' ends in the presence of a labeled α - ^{32}P dCTP and three other dNTPs and the Klenow fragment of DNA polymerase I. P dCTP (specific activity 3000 Ci/mmmole. in the presence of Klenow fragments, was used to fill in the recessed ends of Bgl II and Ava II. Fig. 4.6 shows that Bgl II - Ava II fragment is 1 Kb away from the spacer. The 3' labeled Bgl II - Ava II was digested with Pvu II in a medium strength buffer until completion. The digested fragments were fractionated on ethidium bromide containing low melting agarose (1.0%) prepared in 50 mM Tris-borate and 1 mM EDTA, pH 8.3 for 3 - 4 hours at 100V. Bands containing the digestion fragments 0.8 Kb Bgl II - Pvu II and 0.08 Kb Pvu II - Ava II were removed and transferred to separate Eppendorf tubes, diluted with electrophoresis buffer to obtain 0.2% agarose concentration.

The tubes were heated at 100°C for 15 min. to denature the DNA fragments. Similarly, the uncut Bgl II - Ava II fragment was denatured and used as control. The three probes (3' labeled), Bgl II - Ava II, Bgl II - Pvu II, Pvu II - Ava II were separately hybridized to nitrocellulose filter bound cp rRNAs. Results are presented in Fig. 4.9 and Table 4.8. The 3' labeled, denatured strand of Bgl II - Ava II and Bgl II - Pvu II hybridized to 16S rRNA whereas 3' labeled, denatured strand of Pvu II - Ava II did not hybridize to 16S rRNA. These results clearly indicate that

the direction of transcription is from Bgl II site towards the Pvu II site. The rationale behind this conclusion is that filter bound 16S rRNA will hybridize to the coding strand, labeled or unlabeled. The 3' labeled strand of Bgl II - Pvu II DNA contains the transcription region complementary to 16S rRNA, whereas the 3' labeled strand of Pvu II - Ava II is the RNA-like strand, hence does not hybridize. These findings were further supported by the hybridization of 3' labeled, denatured, strand of subfragments, Ava II - Sma I and Sma I - Ava II of Ava II - Ava II (Fig. 4.9 and Table 4.8). Ava II - Ava II overlaps the 16S rDNA and the spacer. The above evidence in conjunction with 16S rRNA mapping, enables us to predict that 16S rRNA is transcribed from a single DNA strand and is transcribed from Bgl II - Ava II.

4.3.8 Sequence divergence studies

The degree of sequence divergence between soybean and tobacco inverted repeat was obtained by aligning the cleavage map of the single segment of the inverted repeat and calculating the fraction of cleavage sites shared (s) from the equation of Brown et al (1979). i.e. $s = z / (x + y - z)$ where x and y are cleavage sites on the soybean and tobacco inverted repeat region respectively, and z is the number of cleavage sites shared.

The information regarding the cleavage map of tobacco was obtained from the sequence studies of Tohdoh and Sugiura (1981), and Takaiwa and Sugiura (1982a, 1982b). Maps were aligned with respect to the Pvu II site in 16S rDNA.

The minimum number of base substitutions per base (m) is given by the following equation (Brown et al, 1979)

$m = (1 - s)/n$ where n is the number of base pairs per cleavage site. In this case, $n = 6$, as cleavage sites of restriction endonucleases recognizing hexanucleotides were only compared.

The estimated number of base substitutions per base (p) is calculated from the equation derived by Upholt (1977)

$$p = (-\ln s) / n.$$

Two other quantities, 100 m and 100 p are required to calculate minimum and estimated percentage sequence divergence respectively.

Values calculated from the afore-mentioned equations after comparison of the cleavage sites in Fig. 4.10 are collected in Table 4.9. The EcoR I and Sma I sites in the 16SrDNA of soybean and tobacco are considered to be at a homologous position although they are relatively inverted with respect to each other (Fig. 4.10). The cleavage sites of an enzyme on two different DNAs are considered to be at homologous position, if positioned within 1 map unit, 1% of the length of DNA under investigation (Brown et al, 1979). In the present case, one map unit equals 1% of the length of the inverted repeat (23.5 Kb) or 235 bp.

Based upon these results, the non-rDNA region (inverted repeat minus rDNA cistron) has the highest sequence divergence (11.6%), whereas 16S rDNA least divergent. i.e. It is almost similar in base sequence. Notice that the sequence divergence of the entire segment of the inverted repeat is less than that of the non-rDNA region indicating that more base sequence differences are expected to occur in the non-rDNA region as compared to the rDNA region. Among the portions of the rDNA unit compared, 16S - 23S spacer has the highest percentage of sequence divergence (9.7%).

The cleavage sites appear to be randomly distributed throughout the rDNA unit, although there is a notable lack of sites in the spacer region juxtaposed to 23S rDNA. The Pvu II and Xho I sites appear to be highly conserved as earlier pointed out by Palmer and Stein (1982) and Palmer et al, (1983).

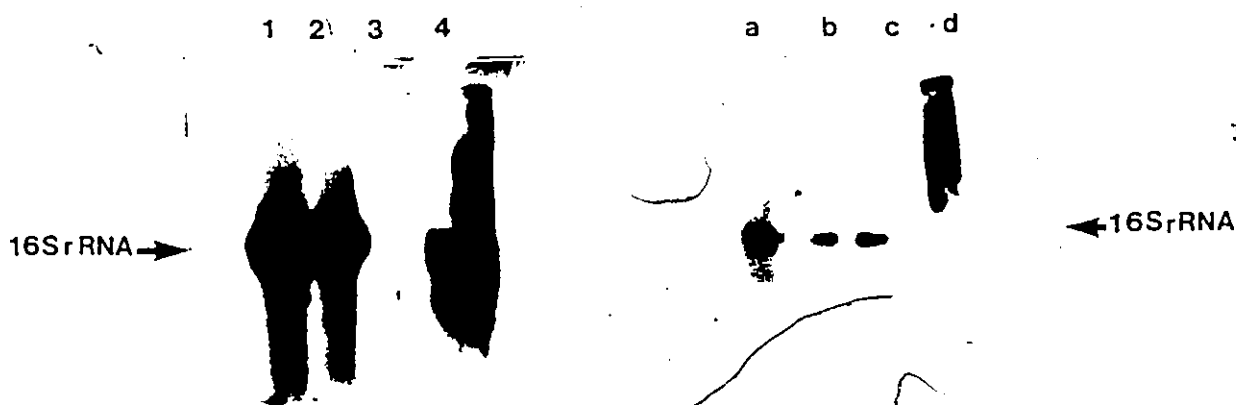


Figure 4.9 Determination of direction of transcription 16S rRNA gene. Ribosomal RNAs from *E. coli* and soybean were extracted, electrophoresed on agarose, transferred to nitrocellulose paper, hybridized to the following 3' labeled, denatured subfragments of rDNA region and autoradiographed.

- (i) 1.0 Kb Bgl II - Ava II fragment to *E. coli*, (lane 1) and soybean rRNAs (lane 2), 0.8 Kb Bgl II - Pvu II fragment (lane 4) and 0.08 Kb Pvu II - Ava II fragment (lane 3) to soybean rRNAs.
- (ii) 1.85 Kb Ava II - Ave II fragment to soybean rRNAs (lane a), 1.0 Kb Ava II - Sma I fragment to *E. coli* rRNAs (lane b) and soybean rRNAs (lane c), 0.8 Kb Sma I - Ava II to soybean rRNAs (lane d).

The size of soybean rRNAs was determined by reference to the mobility of *E. coli* rRNAs.

Table 4.8 Determination of Direction of Transcription. Summary of hybridization of 3' -labeled denatured sub-fragments of the 16S rDNA region to filter bound E. coli and soybean rRNAs.

3' -labeled probe	Filter fixed rRNAs	
	<u>E. coli</u>	Soybean
1.00 Kb Bgl II-Ava II	+	+
0.80 Kb Bgl II-Pvu II	ND	+
0.08 Kb Pvu II-Ava II	ND	-
1.85 Kb Ava II-Ava II	ND	+
1.00 Kb Ava II-Sma I	+	+
0.80 Kb Sma I-Ava II	ND	-

Sizes of fragments are in Kb.

ND - not determined.

Table 49 Quantitative comparison of inverted repeat (one unit only) of soybean and tobacco chloroplast DNA restriction maps

Portion of the Inverted Repeat compared	Molecular Sizes (in Kb)		Restriction Sites compared	Restriction sites shared	Fractions of sites in common	Percentage sequence divergence ^a	
	Soybean	Tobacco ^b				Minimum(M)	Estimated(P)
Inverted Repeat (one unit only)	23.2	23.5	32	23	0.72	4.7	5.5
Inverted Repeat minus rDNA	15.5	16.0	4	2	0.5	8.3	11.6
Entire rDNA	7.2	7.0	28	21	0.75	4.2	4.8
23S rDNA region (4.5-5S)	3.6	3.4	16	13	0.81	3.2	3.5
16S-23S rDNA spacer	2.1	2.1	9	5	0.56	7.3	9.7
16S rDNA	1.5	1.5	3	3	1.0	0.0	0.0

a) Values are calculated as 100 M or 100 P, where M and P are minimum and estimated substitution per base pair respectively, from Brown et al (1979) after comparison of the cleavage maps shown in Fig. 410.

b) The sizes of the inverted repeat, 23S rDNA, 16S-23S rDNA spacer and 16S rDNA are from Fluhr and Edelman (1981a) Takaiwa and Sugiura (1982a), Takaiwa and Sugiura (1982b) and Tohdoh and Sugiura (1981) respectively.

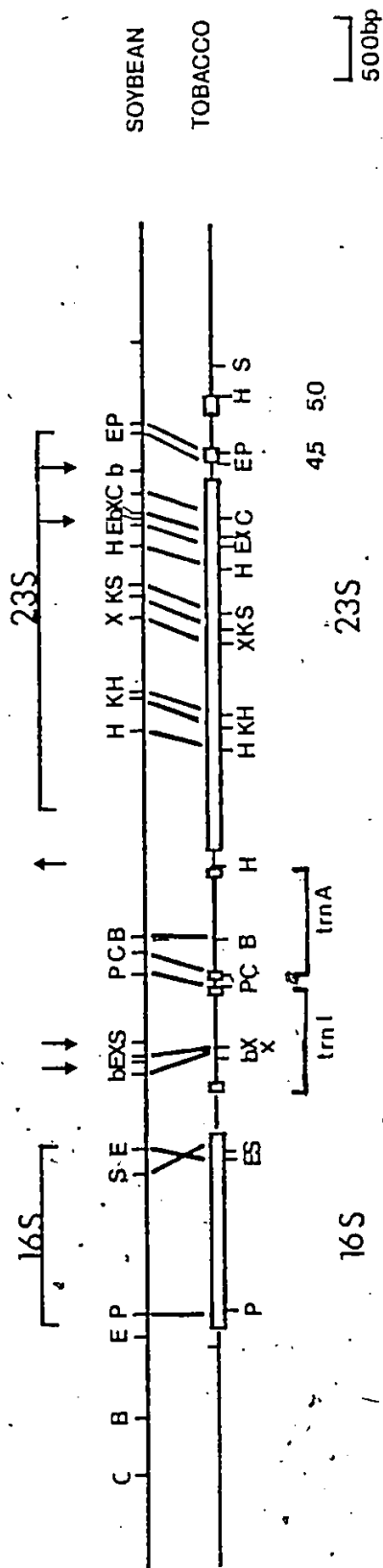


Figure 4.10 Comparison of soybean and tobacco restriction maps of the rRNA cistron. 16S and 23S rRNA are represented by flat brackets (soybean) and open bars (tobacco). Restriction sites are shown for BamH I - b, Bgl II - B, EcoR I - E, Hind III - H, Kpn I - k, Pvu II - P, Sac I - C, Sma - S and Xho I - X. The sites exclusively found in tobacco and soybean are represented by (+) and (-) respectively. The positioning of restriction sites of the tobacco rDNA is based upon Takaiwa and Sugiura (1982a, b), Tohdoh and Sugiura (1982).

4.4 DISCUSSION

4.4.1 Presence of two rRNA cistrons per chloroplast DNA

The hybridization results presented above, suggest that DNA sequences complementary to the 16S and 23S rRNA are present on the soybean chloroplast genome. The labeled rRNA probes hybridized to bimolar fragments (Table 4.1). The positioning of these bimolar fragments implies that rRNA encoding region is present once per segment of the inverted repeat hence, two copies of rDNA gene are present per chloroplast genome. The two sets of the rRNA genes are separated by approximately 20 Kb in one direction and 116 Kb in the other.

4.4.2 Organization of rDNA unit (rrn operon)

Fine mapping of the rDNA unit (rrn operon) suggests that each set contains a single copy of 16S and 23S rRNA genes. The genes are separated by 2.3 - 3.0 Kb spacer. at least 2.3 Kb with an upper limit of 3.0 Kb. Therefore, in each rDNA unit, the rRNA genes are arranged in the sequence 5'-16S - spacer - 23S-3'. The two ribosomal RNA genes are located at the facing end of the inverted repeat (Fig. 4.2), therefore are arranged in inverted orientation with respect to each other. This is clearly established by referring to Pvu II fragments (4.1 Kb and 2.7 Kb) containing rRNA genes (Fig. 4.2, Table 4.1).

4.4.3 rRNA is transcribed from 16S towards 23S

The direction of transcription of 16S rRNA gene was defined by hybridizing 3' labeled ~~strands~~ of 16S rDNA terminal regions to filter bound chloroplast rRNAs. The results indicate that the direction of transcription is from 16S towards 23S. The mode of transcription exhibited by soybean chloroplast genome is similar to that of other plant chloroplast rrn operons examined. Bedbrook et al (1977) have shown that corn chloroplast 16S, 23S and 5S rRNAs are transcribed from the same strand of DNA. Similar observations were made by Rochaix and Malnoe (1978) in Chlamydomonas, and Bohnert et al (1976) in spinach.

It is believed that in plants (Hartley and Ellis, 1973, Hartley, 1979), 16S and 23S are transcribed as a common precursor, whereas a separate transcription signal initiating the synthesis of 5S rRNA has been proposed (Takaiwa and Sugiura, 1980).

4.4.4 Comparison of Soybean rRNA genes with other plants

Physical mapping of rRNA genes of different plant chloroplast genomes has revealed that plants with inverted repeats, contain two sets of rRNA genes as opposed to one in plants such as pea and broadbean which lack the repeat. The overall positioning, organization and copy number of rRNA genes in soybean cpDNA, is comparable to other cpDNAs possessing inverted repeats. The length of the spacer is

different in various cpDNAs examined. The spacer of Euglena rDNA region is much smaller (258 bp) than those of corn and tobacco (2.0 Kb; Koch et al, 1981, Takaiwa and Sugiura, 1982). Yet another striking difference is the split tRNA genes in corn and tobacco in contrast with to uninterrupted tRNA genes in Euglena. The sequence studies in conjunction with other studies (Hartley, 1973, Hartley and Ellis, 1979) suggest that spacer DNA is transcribed as a part of the rRNA operon transcription unit which is subsequently processed to generate rRNA and tRNAs.

4.4.5 Prokaryotic feature of cpDNA

Total soybean rRNA contains species which have electrophoretic mobilities similar to those of E. coli 16S and 23S rRNAs. Further information regarding prokaryotic feature is obtained from hybridization studies. E. coli 16S and 23S rRNA probes gave autoradiogram patterns similar to those of soybean and tobacco rRNAs (Tables 4.1, 4.6 and 4.7). Moreover, 3' labeled denatured strands of soybean 16S rDNA region hybridize equally with soybean and E. coli filter bound 16S rRNA (Fig. 4.9).

These results illustrate that there is sufficient homology between rRNA encoding region of E. coli, soybean and tobacco cpDNA to allow interchangeable use of their 16S or 23S rRNA to be used as probes to identify rRNA genes of the other cpDNAs. The lesser degree of hybridization

observed with E. coli 23S rDNA, under similar reaction conditions might be due to differences in sequence homology but oversimplified interpretation of the results may lead to incorrect interpretations. E. coli 5S rRNA did not hybridize to filter bound soybean cpDNA, indicating a degree of homology too low to warrant use of E. coli 5S rRNA as a probe.

The exact extent of homology between different portions of the E. coli and soybean rDNA unit will be determined only when the latter has been sequenced.


4.4.6 Chloroplast rDNA sequence divergence

The quantitative comparison of cleavage maps and the subsequent calculation of the percentage of sequence divergence are based upon certain assumptions (Brown et al, 1979). These are (a) that each alteration in base sequence is due to the substitution of one base pair and not from rearrangement, insertion or deletion. (b) that the methylation pattern, if any, remains unchanged. (c) that all the base positions in the sequence are equally susceptible to substitution. The validity of these assumptions results in conservative estimation of the sequence divergence (Brown et al, 1979) whereas invalidity of the assumption results in the value of p that underestimates the amount of substitution. (Upholt, 1977; Brown et al, 1979). The results summarized in Table 4.9

show more sequence divergence (11.6%) in the non-rDNA portion of the inverted repeat than in the rDNA region (4.8%). The sequence dissimilarity for the non-rDNA region is moderate, considering the Interfamily level of comparison (Leguminosae vs Solanaceae). It has been shown by such quantitative comparisons (Palmer et al, 1983) and DNA hybridization analysis (Palmer and Thompson, 1982) that there may be as much cpDNA base sequence dissimilarity within Leguminosae as among all the dicotyledonous plants.

From the base sequence divergence values (Table 4.9), fine mapping (Tables 4.4 and 4.5) and hybridization analysis (Tables 4.6 and 4.7) one expects to find high sequence homology between soybean and tobacco 16S rDNA and comparatively lower homology between the 16S - 23S spacer. Moreover, the alignment of the two cleavage maps reveals a downstream drift of the entire cleavage sites of soybean 23S rDNA (Fig. 4.10). This displacement may be due to an addition either in the 16S - 23S spacer or at the start of 23S rDNA.

Sequence analysis of the soybean rDNA unit is essential to reveal its molecular architecture and the extent of homology with tobacco rDNA operon. It will be interesting to find out if the conclusions drawn from the comparison of the cleavage maps will be comparable to those based on sequence analysis.



Chapter V

MAPPING OF CHLOROPLAST tRNA GENES

5.1 INTRODUCTION

In recent years the structure and organization of organelle tRNAs and tRNA genes have been studied in various systems. Several tRNA gene maps have been published recently (Euglena; Hallick, 1983 , Keller et al 1980 , Kuntz et al 1982; spinach Driesel et al, 1979; common bean ; Mubumbila et al, 1983).

Among the plastid genes, those for tRNA are of considerable interest Steinmetz et al (1983) have listed a number of reasons to undertake such studies. These include (a) tRNAs to be used as hybridization probes are easily purified, identified and labeled (Driesel et al 1979). (b) tRNA genes are readily located on the chloroplast genome by molecular hybridization (Driesel et al 1979). (c) the small sizes of the tRNA genes allow rapid sequencing of the transcribed portions as well as the flanking regions that may contain possible regulatory signals. (d) the large number of tRNA genes on the small chloroplast genome facilitates comparisons of a relatively large number of promoter and terminator sequences. (e) tRNA genes are helpful markers in the construction of restriction maps.

(f) comparisons of the arrangements of tRNA genes in cpDNA of various plants may contribute to the understanding of the evolution of chloroplast genomes and (g) Knowledge of tRNA gene numbers and sequences will help to elucidate features of the chloroplast genetic code and of its protein synthesizing system although the final answers will come from sequencing of the mature, functional tRNAs.

The chloroplast tRNA genes are dispersed throughout the genome, with regions of high and low frequency. This dispersion of tRNA genes over the entire chloroplast genome makes it likely that many, or perhaps all, are transcribed independently. Therefore the individually transcribed genes must contain their own promoter regions and sequences involved in transcription termination.

Construction of restriction endonuclease cleavage site map of chloroplast DNA and the purification of individual chloroplast tRNAs by two-dimensional gel electrophoresis, in recent years, has allowed considerable progress in the localization of tRNA genes on the chloroplast genome. The mapping of tRNA genes has been reported on Euglena (Hallick, 1983; Keller et al, 1980; Kuntz et al, 1982) ; Chlamydomonas (Malnoe and Rochaix, 1978) Spinach (Driesel et al, 1979) and bean (Mubumbila et al, 1983) cpDNAs. Total cp tRNAs were used to map tRNA genes in Euglena and Chlamydomonas whereas specific tRNAs were utilized to position their respective genes in spinach and bean.

So far, all the above studies have shown that chloroplast isoaccepting tRNAs are often coded for by different genes located in different areas of the chromosome. Further studies have led to the presence of tRNA^{Ile} and tRNA^{Ala} genes in the spacer between 16S and 23S rDNA in Chlamydomonas (Malnoe and Rochaix, 1978), spinach (Bohnert et al, 1979), Euglena (Keller, et al 1980, Orozco et al, 1980, Graf et al, 1980) common bean (Mubumbila et al, 1983) and Zea mays (Koch et al, 1981). Similar situation prevails in E. coli (Young et al, 1977).

In an effort to understand the features of the chloroplast protein synthesizing system and its genetic code, a number of chloroplast tRNA genes have been sequenced as in maize (Schwarz et al 1981, Koch et al 1981, Steinmetz et al 1982), spinach (Kashdan and Dudock 1982), tobacco (Kato et al 1981 and Tohdoh et al 1981) and Euglena (Graf et al 1980 and Orozco and Hallick 1982). These studies show features of nuclear tRNA genes (3' terminal CCA not encoded) and prokaryotic tRNAs (high sequence homology) and some unique features peculiar to chloroplast tRNAs (presence of very long introns). In maize, the chloroplast tRNA^{Ala} gene has an intron of 806 bp and that of tRNA^{Ile} an intron of 949 bp, both located between the second and third base after the anticodon. However, in the case of Euglena, the ribosomal spacer is only 259 bp long. Most of the sequence studies on chloroplast tRNA genes so far reported in Euglena and maize,

have shown that the 3' CCA end is not coded for and is therefore added post-transcriptionally.

We have undertaken this project to map a few tRNA genes of a commercially important plant, the soybean and to compare the results with existing information on spinach and common bean.

5.1.1 Isolation of chloroplast tRNA

The chloroplast tRNAs were prepared from 10-12 day old cotyledons according to Burkard et al (1970). The chloroplast pellet obtained upon centrifugation of a liter of the homogenate was dissolved in a minimum volume (5 ml) of 10 mM Tris HCl, pH 7.4, 10 mM MgCl₂, 1% SDS with the help of a fine brush and extracted against watersaturated phenol (8:2 v/v) by constant stirring at 4°C for 2 hrs. The aqueous phase was separated by centrifugation at 4000 g for 15 min, recovered, made 200 mM with respect to potassium acetate, mixed with two volumes of 95% ethanol and stored at -20°C for 8-10 hrs. The RNA precipitate was obtained by centrifugation, dried, dissolved in 10 mM Tris-HCl, pH 7.5, 200 mM NaCl, 1 mM MgCl and incubated with DNase I (Sigma DN-100) at 4°C for 90 min. The DNase treated solution was loaded onto a DEAEcellulose column (3 ml bed volume / kg of starting material), pre-equilibrated with 50 mM Tris HCl, pH 7.4, 200 mM NaCl, and washed with the same buffer until 260 dropped below 0.02 units. The tRNAs were eluted with

1.0 M NaCl, collected, de-aminoacylated and precipitated with two volumes of cold ethanol. The tRNA pellet was either used for separation of the various species or lyophilized and stored at -20°C .

5.1.2 Two-dimensional polyacrylamide gel electrophoresis

The total cp tRNA population was fractionated on two-dimensional polyacrylamide gel, (PAG), in order to recover the individual tRNA species (Swamy and Pillay, 1982). The first dimension was conducted at 4 C in 10% vertical polyacrylamide gel prepared in 10% acrylamide, 0.4% N-N'-methylene bis acrylamide (BIS), 4 M urea in 0.1 M tris-borate buffer, pH 8.3, 4 mM EDTA for 36-40 hrs. at 450 volts. The acrylamide gel was polymerised with 0.001 % ammonium persulphate and 1 ul/ml of TEMED. A narrow strip of the PAG containing the bands was cut, and placed horizontally between two glass plates. Polyacrylamide solution containing 20% acrylamide, 0.8% BIS, 4 M urea in 0.1 M Tris borate buffer pH 8.3 was added. The second dimension gel was electrophoresed at 350 volts for 140 hrs. The tRNA spots were visualized by methylene blue staining.

5.1.3 Separation and identification of individual spots

Individual tRNA spots were cut out of the gel, transferred to 1.5 ml microfuge tubes, crushed and mixed with equal amount of phenol saturated against high salt Tris

EDTA (500 mM NaCl, 10 mM Tris HCl, pH 8.0, 1 mM EDTA) for 6-8 hrs at room temperature. The aqueous phase was separated by centrifugation and mixed with two volumes 95% ethanol to precipitate tRNA. 5 ug/ul of rRNA was used as carrier. The purified tRNAs was identified by aminoacylation, using the E. coli synthetase and 3 Hamino acids.

5.1.4 Labeling of tRNA

The identified tRNA was treated with snake ~~venom~~ phosphodiesterase to remove the terminal adenosine of the -CCA-3 end, and re-extracted with phenol to denature the phosphodiesterase and the tRNA recovered by Lyophilization. The tRNAs were enzymatically labeled using ($\alpha^{32}\text{P}$) ATP, CTP, yeast nucleotidyl transferase (Rether et al, 1974). The radiolabeled tRNAs were separated from the unincorporated label by RPC-5 column chromatography. Only fractions containing the highest radioactivity were used as probes. The probe was denatured in 50% formamide before hybridization.

5.1.5 tRNA DNA hybridization

Nitrocellulose filters were soaked in hybridization solution containing 50% formamide, 2XSSC for 10 min. The filters were then placed in Dazey heat-sealable bags and fresh hybridization solution was added along with the probe.

The bags were sealed and incubated at 37°C for 10-12 hrs. The filters were washed as described earlier (Chapter 4).

5.2 RESULTS AND DISCUSSION

5.2.1 The three tRNAs have separate chloroplast genes

Results from our laboratory (Pillay and Cherry 1974, Sinclair and Pillay 1981) have consistently shown the existence of only two chloroplastic tRNAs^{Leu} compared to three distinct tRNAs^{Leu} in Phaseolus vulgaris (Osorio Almeida et al, 1981) and spinach (Steinmetz et al, 1979). It was of interest to undertake some rigorous tests to resolve the number of tRNAs in the soybean system. Using a combination of column chromatography and 2-dimensional PAGE it was shown that there are, indeed three chloroplastic tRNAs^{Leu} (Guillemaut - personal communication). The next step was to determine if these three different tRNAs^{Leu} were coded for by three separate chloroplast genes.

The results (Table 5.1, Fig. 5.1) clearly indicate that the labeled tRNA^{Leu}_{CmAA} tRNA^{Leu}_{UmAA} tRNA^{Leu}_{UA^mG} hybridized to inverted repeat, to large and to small single copy regions respectively. The gene for tRNA^{Leu}_{UmAA} (trnL1) is located in 5.6 Kb Xho I, 12.5 Kb downstream from the inverted repeat, trnL2 is represented once per unit of the inverted region as tRNA^{Leu}_{CmAA} hybridized to 23.2 , 17.8 Kb Kpn I fragments (Fig.5.1) trnL3 is positioned in the 9.6 Kb Pvu II fragment in the small single copy region.

The overall localization of the three tRNAs^{Leu} genes in soybean is similar to the results reported in another legume, common bean (Mubumbila et al, 1983) but are different from spinach (Driesel et al, 1979, Steinmetz et al, 1979). The gross differences are in the positioning of the trnL1 and trnL3. In comparison to soybean, spinach trnL1 gene (gene sets) is relatively more remote from trnL2 and trnL3, is present twice instead of once (Fig. 4 in Mubumbila et al, 1983 and Fig. 6 in Driesel et al, 1979). The sequence rearrangement may be consequence of evolutionary change as suggested by Palmer and Thompson (1982). The functional implications of such structural rearrangements on the expression of the plastid genes are yet to be determined. Presence of the third species of tRNAs^{Leu}_{CmAA} was undetected due to the lower amount of tRNA. There is considerably difference in the extractable amounts of three tRNAs (Guillemaut-personal communication). Surprisingly, hybridization data shows that the tRNA^{Leu}_{CmAA} is present in duplicate sets. The lower availability tRNA might be due to some sort of control at the transcriptional or post-transcriptional level. It will be of interest to study the expression of tRNAs. to find out if such a control exists in vivo or in vitro. In addition to the above three tRNA genes we have localized the genes for tRNA^{Ser} and tRNA^{Phe}. These two tRNAs were isolated, purified, labeled and hybridized as in the case of leucine.

Table 5.2 shows that the genes for both tRNA^{Ser} (trnS) and tRNA^{Phe} (trnF) are positioned in the large single copy region. The trnS gene maps in the 5.6 Kb Sac I and trn F in 23.2 Kb Sac I, the localization of these tRNA genes is comparable to that of the common bean (Mubumbila, 1983).

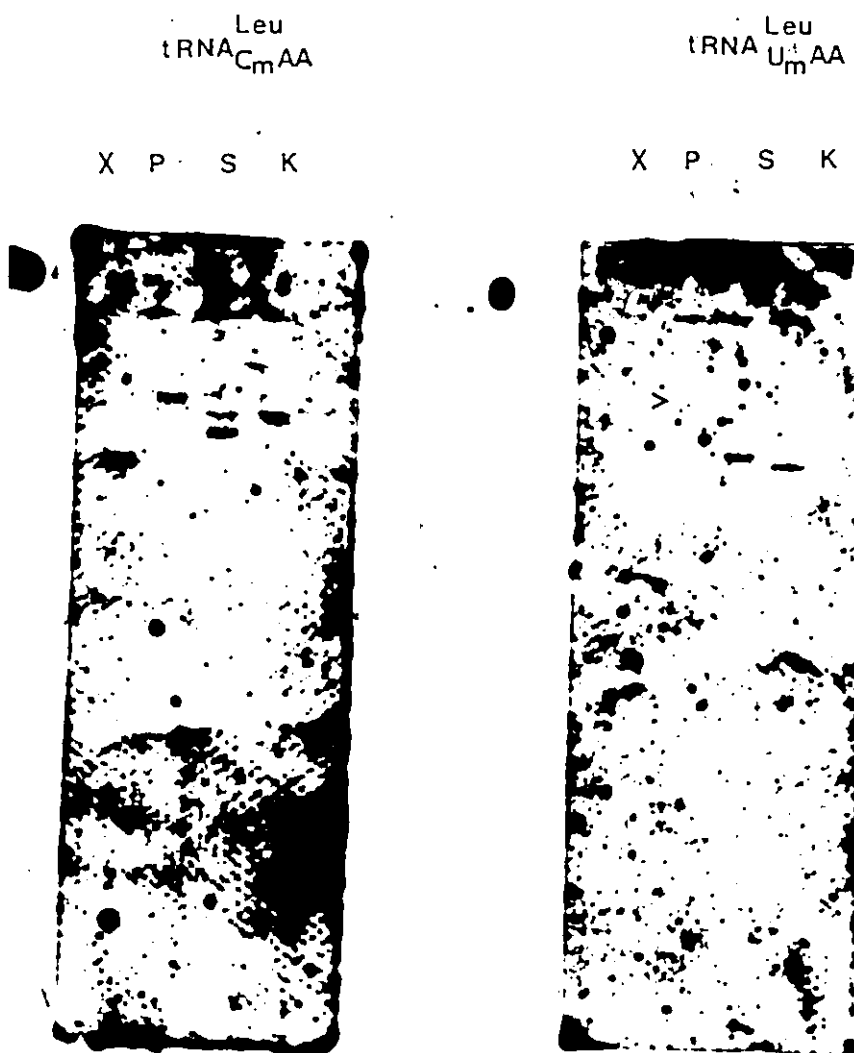


Figure 5.1 Hybridization of ^{32}P labeled tRNA^{Leu}_{CmAA} and tRNA^{Leu}_{UmAA} to filter-bound digests of Kpu I (K), Pvu II (P), Sac I (S) and Xho I (X). The results of hybridization are summarized in Table 5.1.

Table 5.1 Localization of soybean chloroplast tRNA^{Leu} genes by hybridization of ³²P labeled tRNAs^{Leu} to filter bound soybean cpDNA restriction digests.

Restriction digests of	tRNA ^{Leu} UmAA	tRNA ^{Leu} CmAA	tRNA ^{Leu} UA ^{m7} G
Kpn I	12.0	27.5, 24.0	24.0
Pvu II	34.0	41.0, 34.0	9.6
Sac I	13.6	23.2, 17.8	19.8
Xho I	5.6	12.0	23.4

Sizes of the fragments are in Kb.

Table 5.2 Localization of Serine and phenylalanine tRNAs by hybridization of $\alpha^{32}\text{P}$ ATP labeled probes to the filter fixed soybean cpDNA fragments.

Restriction digests	Soybean tRNAs	
	tRNA ^{Ser}	tRNA ^{Phe}
Kpn I	8.4	27.5
Pvu II	8.6	34.0
Sac I	5.6	23.2
Xho I	-	-

Sizes of fragments are in Kb.

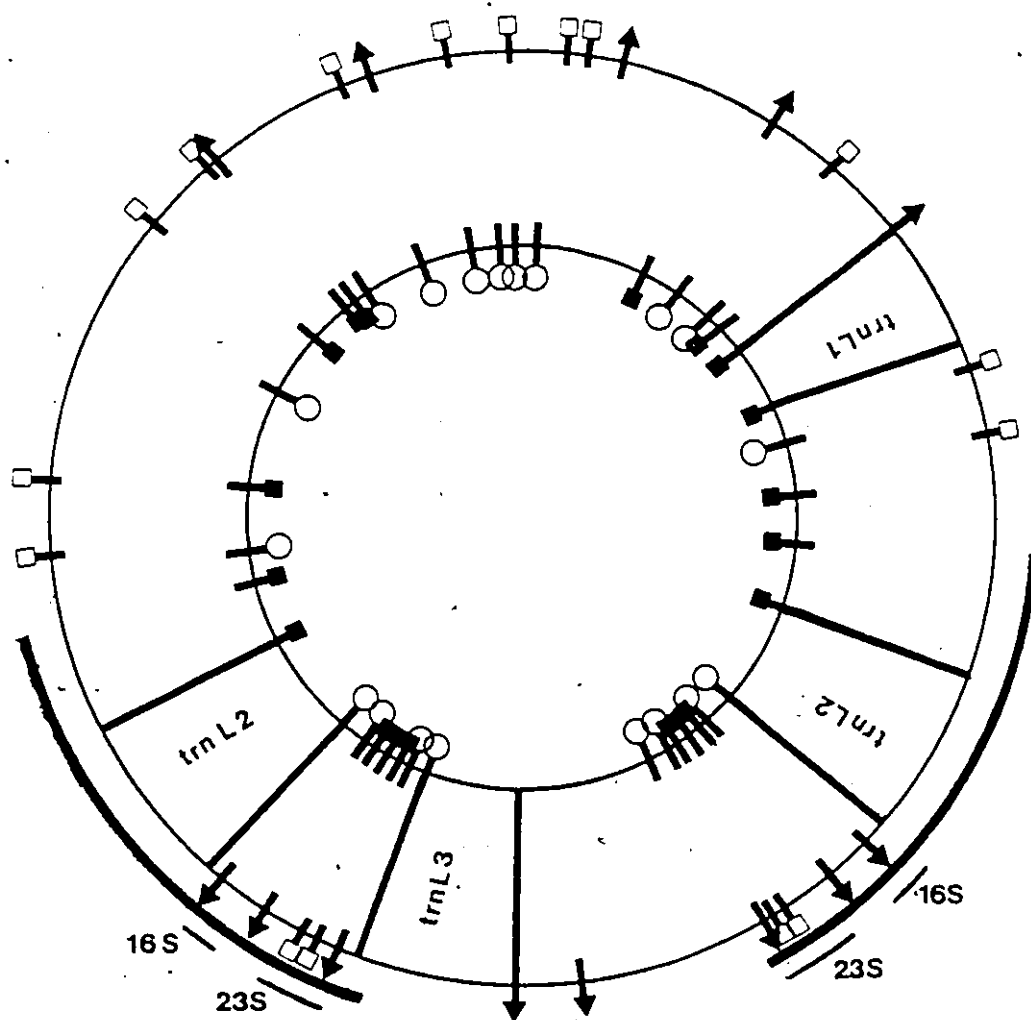


Figure 5.2 Localization of the trn L1, trn L2 and trn L3 genes on the physical map of soybean cpDNA. The cleavage sites of the following restriction enzymes are mentioned. Kpn I (\square), Pvu II (\blacktriangle), Sac I (\circ) and Xho I (\blacksquare).

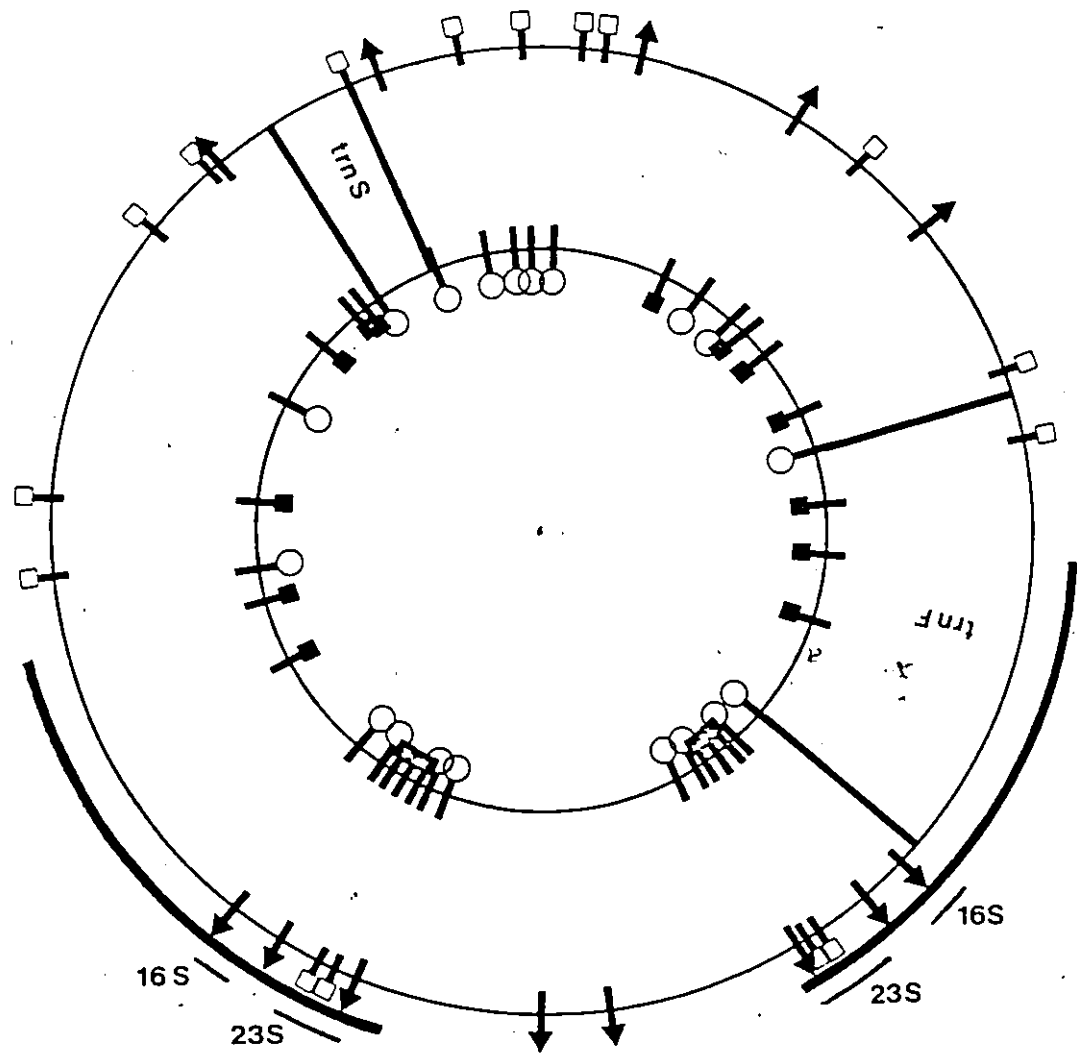


Figure 5.3 Positioning of *trnS* and *trnF* on the physical map of soybean cpDNA. The cleavage sites of the following restriction enzymes are shown.
Kpn I (□), Pvu II (▲), Sac I (○) and Xho I (■).

Chapter VI

POSITIONING OF PLASTIDS PROTEIN ENCODING GENES

6.1 INTRODUCTION

The number and diversity of structural and functional proteins present in or associated with fully grown chloroplast is considerably higher than the theoretical coding capacity of the chloroplast genome (Ciferri, 1978). It is evident from the selective inhibitor analysis (Ellis, 1977), in vitro labeling (Ellis, 1978), and studies of heat-treated plants (Feierabend and Schrader-Reichhardt, 1976, Feierabend and Wildner, 1978), that the majority of the chloroplast proteins/polypeptides are encoded by the nuclear genome, synthesized on cytoplasmic ribosomes and then transported across the chloroplast envelope by some post-translational mechanism (Dobberstein et al, 1977, Grossman et al, 1980).

Although a typical chloroplast DNA (120 Kb - 150 Kb) could possibly encode for 100-200 average sized (50,000) polypeptides, yet only few cpDNA region encoding plastid polypeptides have so far been identified and mapped. These include the large subunit for ribulose-1,5- biphosphate carboxylase-oxygenase, 8 Kd subunit (III) of CF_o, plus three subunits (α , β , ϵ) of the coupling factor CF₁, components

of ATPsynthase (ATPase), elongation factor T and G of plastid protein synthesis, cytochrome f, 15 Kd protein of cyt bf and 32 Kd thylakoid membrane protein.

6.1.1 Gene for Large Subunit of RUBISCO (rbcL)

Ribulose-1,5-bisphosphate carboxylase oxygenase (RUBISCO), one of the distinguished proteins of plants, is responsible for CO₂ fixation and oxidation of ribulose-1,5-bisphosphate. RUBISCO is composed of eight large subunit (LS) of nearly 55,000 molecular weight and eight small subunits (SS) of approximately 15,000 molecular weight (Lorimer, 1981). The large subunit is cpDNA coded (Chan and Wildman, 1972, Coen et al, 1978) and ~~contains~~ the catalytic site. The small subunit is nuclear DNA coded, and transported across chloroplast envelope by post translation mechanism (Chua and Schmidt, 1978).

The gene for large subunit of RUBISCO (rbcL) has been mapped on the chloroplast genomes of several plants including *Chlamydomonas reinhardtii* (Malnoe et al, 1979), *Euglena gracilis* (Steiglier, 1982), tobacco (Seyer et al, 1981), spinach (Whitfield and Bottomley, 1980), corn (Coen et al, 1977), broad bean (Ko, 1982). *Osmunda cinnamomia* (Palmer and Stein, 1982).

In plants containing inverted repeat, the rbcL maps in the large single copy region either just outside the inverted repeat (mung bean, common bean, soybean) or at a

position relatively more remote (corn, tobacco). Steigler et al (1981) localized the position of *rbcl* in Euglena by hybridizing restriction fragments containing internal sequences of Chlamydomonas and corn, and have reported the presence of 0.5- 1.1 Kb intervening sequence. The *rbcl* gene from corn (McIntosh et al, 1980, Poulsen, 1981), spinach (Zurawski et al, 1981), Chlamydomonas (Dron et al, 1982), and cyanobacterium, Anabaena 7120 (Curtis and Haselkorn, 1983) has been sequenced. The corn and spinach *rbcl* genes are 84% homologous to each other and are approximately 75 - 77% similar to Chlamydomonas *rbcl* gene sequence. The coding region of Anabaena large subunit of RUBISCO is 71-76% homologous to the analogous portion of corn, spinach and Chlamydomonas.

6.1.2 Genes for subunits of ATP Synthase

ATPase is a multimeric component of thalokoid membrane facing the stroma, and is involved in the ATP synthesis coupled with proton transport. The isolated active ATPase molecule is composed of two parts: the membrane anchored, integral portion CF_o and the external headpiece, CF₁. The membrane anchored, CF_o is formed of 18, 16 and 8 Kd subunits referred to as I, II and III respectively. The projected CF₁, the catalytically active part of the enzyme consists of 59, 52, 37, 19.5 and 16 Kd subunits denoted as α , β , γ , δ and ϵ respectively. Two of the CF_o subunits (I and III) and

three of the CF_1 subunits (α , β and ϵ) have been shown to be products of chloroplast protein synthesis (Mendolia-Morgenthaler et al, 1976, Nelson et al, 1980). It has been adequately defined that α , β and ϵ units are translated from cpRNA; while δ and γ are from polyadenylated RNA (Westhoff et al, 1981). Genes for all the three chloroplast DNA coded (atpA, atpB and atpE respectively) have been positioned in spinach (Westhoff, 1981). The nucleotide sequences of atpB and atpE have been determined in corn (Krebbers et al, 1982), spinach (Zurawski et al, 1982), and pea (Zurawski et al, 1983). Presence of cpDNA sequences complementary to atpA and atpB in *Osmunda* (Palmer and Stein, 1982), and atpB and atpE in wheat (Howe et al, 1982b) have been confirmed by probe hybridization.

In comparison to atpB and atpE, atpA is located relatively more distal, approximately 40 Kb away from rbcL in the cpDNAs examined so far. Genes for β and ϵ subunits of ATPase are separated from the rbcL gene by 350 bp in corn (Krebbers et al, 1982, Zurawski et al, 1982) and have been shown to be cotranscribed as dicistronic message in spinach (Zurawski et al, 1982). The rbcL and ATPase genes are read in the opposite direction with respect to each other.

Comparison of amino acid sequence of β and ϵ subunits of spinach ATPase with *E. coli* ATPase counterparts reveals that β subunit is more conserved (67% homology) as compared to ϵ subunit (26% homology) (Zurawski et al, 1982).

ATPase activity is inhibited by NN'Dicyclohexyl carbodimide (DCCD), a lipid soluble compound that covalently binds to the 8.0 Kd, proton translocating subunit of the CFo component. This DCCD binding polypeptide is coded for and transcribed in the chloroplast (Howe et al, 1982a). In wheat, the gene for DCCD-binding polypeptide (atpH) maps 15 Kb away from rbcL and atpB-atpE complex and does not show any introns. The amino acid sequence predicted from the nucleotide sequence of wheat atpH gene (Howe et al, 1982a), is identical to that obtained for spinach DCCD binding polypeptide by direct amino acid sequence (Sebald et al, 1980).

6.1.3 32 Kd Thylakoid Protein

This protein, unlike the carboxylase and ATPase, is completely encoded, translated, and processed within the chloroplast (Eaglesham and Ellis 1974, Bedbrook et al, 1978). This protein has been recently identified as the shield protein regulating photosystem II electron flow in the thylakoid membrane (Matto et al, 1981) and is responsible for the binding of herbicides such as azodoatrazine (Pfister et al, 1981). The 32 Kd thylakoid protein encoding region (psbA) has been mapped in several plant species including Atriplex (Palmer, 1982), Chlamydomonas (Rochaix, 1981), pea (Palmer and Thompson, 1981), Osmunda (Palmer and Stein, 1982), spinach (Driesel et

al, 1980, and corn (Bedbrook et al, 1978): The highly conserved plastid genes for this membrane protein almost always maps in the species studied, maps near the inverted repeat in the large single copy region. The gene (psbA) has been sequenced in spinach and Nicotiana debneyi (Zurawski et al, 1982).

The nucleotide sequences of psbA gene from spinach and Nicotiana debneyi have been determined to be over 95% homologous (Zurawski, 1982). The size estimation of the protein from the nucleotide sequence reveals a molecular weight of approximately 39 Kd, much larger than that determined by SDS-gel electrophoresis.

Whether the discrepancy between the calculated and estimated molecular sizes is due to processing by some post translational step, or to anomalous behaviour in electrophoretic mobility, as has been reported in other hydrophobic proteins remains to be determined (Darley-Usmar and Fuller, 1981). The deduced amino acid sequence shows complete absence of lysine residues but abundant and clustered of hydrophobic amino acids (Zurawski et al, 1982).

The nucleotide sequences of all the chloroplast coded genes show that codon usage is universal. However, not all the codons are utilized because there is a bias toward adenine or thymidine residues in the third position which might be related to high A-T content of cpDNAs (Whitfield and Bottomley, 1983).

Some of the chloroplast genes exhibit the presence of putative promotor regions comparable to the E. coli "-10" and "-35" consensous sequences and the presence of termination signals, reminiscent of E. coli terminator sequences, in the 3' untranslated region of the gene. Moreover, it has been recently shown that rbcL is expressed in vivo in E. coli by employing either the actual promotor of rbcL or the lambda promotor P. (Gatenby et al, 1982). These observations suggest a prokaryotic nature of control at both transcription and translation levels.

As a prelude to a study of the organization and expression of polypeptide encoding genes, we decided to localize the major plastid protein genes on soybean cpDNA by probe hybridization. Due to high degree of similarity between chloroplast genes from different species, an identified gene from one species may be utilized to position the same gene in other plant species. This approach has been used to position rbcL, atpA, atpB, atpE, atpH, cytF and psbA genes on the soybean chloroplast genome.

6.2 MATERIALS AND METHODS

6.2.1 Materials

Isopropyl thio- β -galactoside (IPTG), 5-bromo-4-chloro-3-indolyl- β -D-galactoside (Xgal), 15-base primer DNA (single stranded) and deoxynucleotides were from Bethesda Research Laboratories, Inc. (Maryland). Radioactive label

was supplied by Amersham (Ontario) and Klenow fragment of E. coli DNA polymerase I was from P. L. Biochemicals.

Media YT Medium ; It consisted of 8g of Bacto tryptone, 5g of Bacto yeast extract and 5g of NaCl per liter of water. 2XYT was twice the concentration of ingredients. To make solid YT plates, 13.0g of bacto agar (Difco) was added to the media before autoclaving. For soft agarose, 6.0g of bacto agar per liter was added, instead. Bacterial strains JM103 lac^{pro} , thi, strA, endA, sbcB15, hsdR4 upE, F' traD36, proAB, lacI^q, Z M15.

6.2.2 Methods

The procedures used for preparation of chloroplast and plasmid DNA, restriction digestion, nick translation, southern filters, hybridization, and washing, have been described in previous chapters.

The following methods are after Messing (1978, 1980).

6.2.2.1 M13 transformation

M13 mp7 is a cloning vector derived from the wild type single stranded M13 phage by insertion of the promoter and the front portion of the β galactosidase gene (β -gal). Furthermore, a polylinker containing the restriction sites of various endonucleases was introduced in the β -gal gene. Intact M13 mp7 synthesizes a functional β -galactosidase forming blue plaques, in presence of IPTG and Xgal.

Inactivation of the β -gal activity by insertion of foreign DNA in the M13 mp7 causes colorless plaques. E. coli K12 strain JM103 (Messing et al, 1981) were

transformed by single stranded M13 DNA. Approximately 50 ml of YT medium was inoculated with a loopful of JM103. cells from minimal agar plate. The growth of the cells was stopped at O.D.₆₀₀ = 0.7 by chilling the cells at 4°C. About 20 ml of the bacterial culture was withdrawn, centrifuged at 6,000 rpm for 5 min. at 4°C. From the remainder of the culture, 0.5ml was used to inoculate a fresh YT medium which served as batch for exponentially growing cells at the time of plating (see below). The bacterial cell pellet obtained upon spinning was suspended in 10 ml of pre-chilled 50 mM CaCl₂ for 20 min. The suspension was re-centrifuged at the same speed and the pellet was resuspended in 2 ml of 50 mM CaCl₂. About 300 μ l of competent cells were added to siliconized 13 x 100 mm tubes containing 0.5 μ g of single stranded DNA and stored on ice for 40 min. The bacteria were subjected to heat shock for 2 min. at 42°C. To this 10 μ l IPTG (100 mM) 0.50 μ l X gal (2% in Dimethyl formamide), 200 μ l of fresh exponentially growing JM103 cells and 3 ml of soft agar at 40°C was added, mixed and plated on fresh YT plates. Upon solidifying, the plates were transferred to 37°C. Turbid plaques were observed in 4 - 6 hours and color reaction in another 2 hours.

6.2.2.2 Preparation of single stranded DNA

Nearly 20 ml of YT medium inoculated with loopful of non-infected JM103 from minimal agar plate and incubated at 37°C for 3 hours on a shaker. The individual colorless turbid plaques were transferred to some of fresh YT medium inoculated with 0.2 ml of 3 hr. old non-infected JM103 culture and incubated at 37°C with constant shaking for 8 - 10 hours. The cells were harvested by centrifugation at 6,000 rpm for 5 min. at 4°C. The pellet was saved for the preparation of replicative form (double stranded circles of M13) and to the supernatant, 2 ml of 40% polyethylene glycol (PEG-6000, Sigma) in 5 M NaCl was added, thoroughly mixed, and centrifuged at 15,000 rpm for 20 min. at 4°C to obtain the nucleic acid pellet. The pellet was dissolved in 10 mM Tris, 1 mM EDTA, pH8.0, extracted twice with phenol, twice with chloroform. To the aqueous phase, 200 µl of 3 M sodium acetate, pH5.0, 2 volumes of ethanol were added and the contents were stored at -80 for 60 to 90 min. The DNA was recovered by centrifugation at 12,500 rpm for 15 min. at 4°C and resuspended in 10 mM Tris and 1 mM EDTA, pH8.0. The preparation was checked for purity on 2% agarose.

6.2.2.3 In vitro conversion of single stranded DNA into Replicative Form.

About 10 µg of 15 nucleotide primer DNA was added to 500 µg of template DNA in 20 mM Tris HCl, pH7.5, 10 mM MgCl₂, 50 mM NaCl, 0.5 mM DTT at temperature conditions which

ensured proper annealing. The template-primer complex was mixed with 10 μCi α - ^{32}P dCTP (specific activity 3000 Ci/mmol), 66 μM each of dATP, dCTP, dTTP, 0.5 units of Klenow fragment of *E. coli* DNA polymerase I and incubated initially at 40 °C for 30 min. and subsequently for another 30 min at room temperature. To this mixture, 5 mM each of d GTP, d ATP, d CTP, d TTP, 2.5 mM ATP, 10 units of T4 DNA ligase was added and reaction continued for 5 hours at room temperature to allow formation of closed circles. The reaction was terminated by heat inactivation. Over 2.0×10^6 cpm / μg of single stranded of DNA was recorded.

6.2.2.4 Preparation of Replicative form

The replicative form of M13 DNA was isolated by lysing the infected bacterial cell pellet obtained upon centrifugation according to procedure modified after Godson and Vapnek (1973).

6.3.1 Gene for large subunit of RUBISCO (rbcL)

The rbcL gene was positioned on soybean chloroplast DNA by either hybridizing 1750 bp EcoRI bounded fragment or 2200 bp Ava I - Bam HI fragment encompassing the spinach carboxylase gene (Erion et al, 1981, Zurawski et al, 1981) to the filter bound Kpn I, Pvu II, Sac I and Xho I digests. Autoradiogram pattern (Fig. 6.1 and Table 6.1) revealed that

both nick translated probes hybridized to 27.5, 3.6 Kb Kpn I, 34.0 Kb Pvu II, 23.2 Kb Sac I and 8.0 Kb Xho I fragments present in the single large copy region just exterior of the inverted repeat (Fig. 6.6) thereby indicating the presence of a region complementary to spinach *rbcL* gene.

6.3.2 Genes for subunits of ATPase

Alpha subunit of CF₁ component (*atpA*)

The nick translated 2.4 Kb spinach Sal I cloned fragment containing the *atpA* gene (Westhoff et al, 1981) was employed to localize the analogous gene on the soybean genome. Fig. 6.2 and Table 6.2 show that it hybridized to 8.4 Kb Kpn I, 8.6 Kb Pvu II, 5.6 Kb Sac I and 27.0 Kb Xho I in the large single copy region, approximately 41.0 Kb distal from the *rbcL* gene indicating the presence of region homologous to spinach *atpA* gene. Beta Subunit of CF₁ (*atpB*):

The region possibly encoding for the beta subunit of ATPase was identified by hybridizing the ³²P labeled 1950 bp EcoRI cloned spinach fragment (Zurawski, 1981) to the blots containing soybean fragments. The results (Fig. 6.3 and Table 6.2) depict that 3.6 Kpn I is the smallest fragment to incorporate the region similar to the spinach *atpB* gene sequences as the adjacently lying 27.5, 12.0 Kb Kpn I did not hybridize to the probe (Fig. 6.6).

Epsilon Subunit of CF (*atpE*)

Single stranded M13mp7

carrying a Sau 3A fragment containing the 3' position of the wheat atp gene was used as template in vitro synthesis of P - dCTP labeled replicative form probe, which was subsequently denatured and hybridized to the soybean cpDNA digests. Except for 23.2 Kb Sac I, the results (Fig.6.5 and Table 6.2) are identical to the atpB gene probe hybridization, thereby indicating that the soybean cpDNA region complementary to spinach atpE gene is completely contained in 3.6 Kb Kpn I.

Subunit III of CF component of ATPase (atpH)

The probe for atpH gene or DCCD binding polypeptide gene was 140 nucleotides Hae III - Hind III fragment of wheat cloned in single stranded M13 mp 7. The fragment was converted into labeled double stranded (RF) in the presence of α -³²P dCTP and Klenow fragment of E. coli DNA polymerase I, denatured and hybridized to soybean cpDNA restriction pattern of Kpn I, Pvu II, Sac I, and Xho I. From the data (Fig.6.5 Table 6.2) the analogous region in soybean could be positioned in 4.2 Kb Sac I, proximal to the 5.6 Kb Sac I fragment containing atpA gene-like sequences.

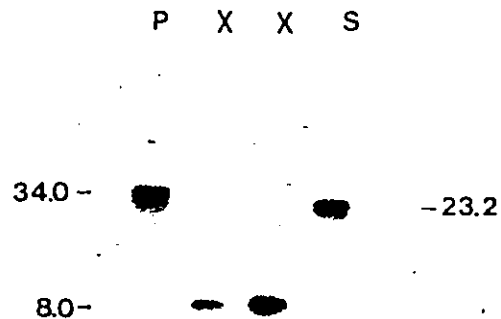


Figure 6.1 Hybridization of Nick-translated 2,200 bp Ava I - BamH I spinach cpDNA fragments, containing *rbcl* gene to nitrocellulose strips containing Pvu II (P), Sac I (S), and Xho I (X) digests. The sizes of the fragments hybridizing to the probe are along the sides.

Table 6.1 Localization of *rbcl* and *psbA* genes by hybridization of ³²P labeled probes to the nitrocellulose filter bound soybean cpDNA.

Single digest	<i>rbcl</i> gene probe (1750bp/2200bp)	<i>psbA</i> gene probe
Kpn I	27.5, 3.6	*
Pvu II	34.0	34.0
Sac I	23.2	23.2
Xho I	8.0	4.5

* - Kpn I digest was not included.

Sizes of fragments are in Kb.

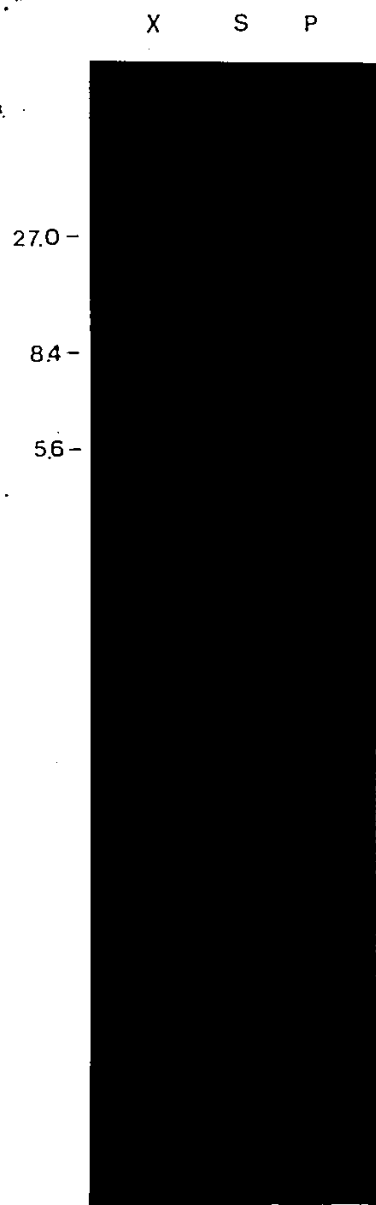


Figure 6.2 Localization of *atpA* gene on soybean chloroplast DNA by Southern hybridization. The 2.4 Kb probe from spinach hybridized to 8.4 Kb Pvu II (P), 5.6 Kb Sac I (S) and 27.0 Kb Xho I. Contact print to the autoradiogram was developed to highlight the bands.

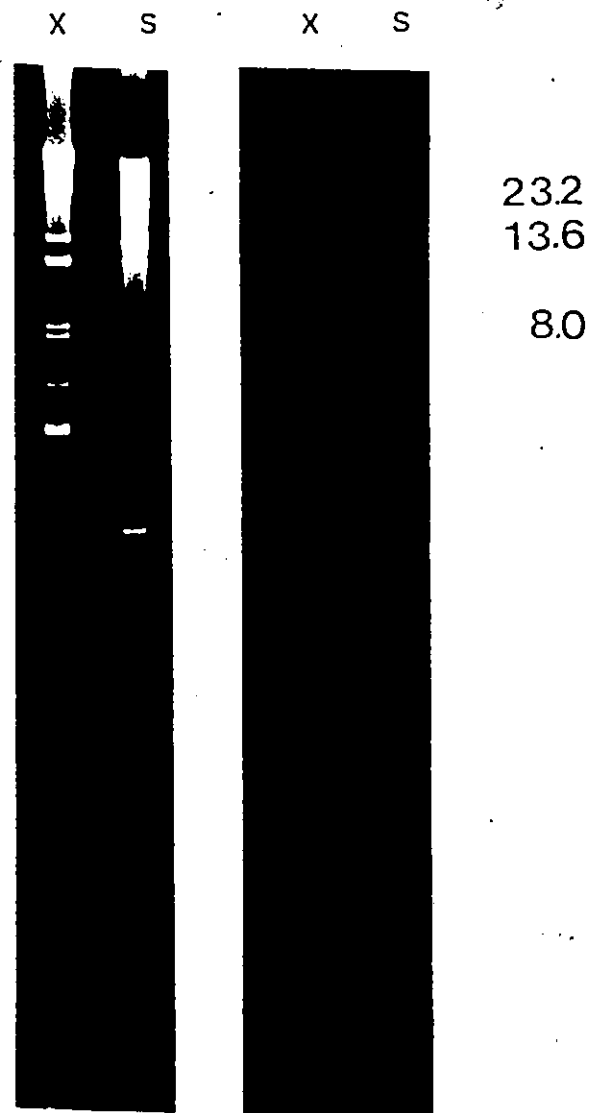


Figure 6.3 Positioning of *atpB* gene on soybean cpDNA by Southern hybridization. The 1950 bp *EcoR* I spinach fragment was Nick-translated, hybridized to filter-bound *Sac* I (S) and *Xho* I (X) fragments of soybean cpDNA and autoradiographed. The sizes of the hybridizing fragment are in Kb.

6.3.3 32 Kd Thylakoid protein gene (psb A)

cpDNA insert containing the psbA gene was labeled by nick translation procedure according to Maniatis et. al (1975) and hybridized to single digests of soybean cpDNA. The hybridization data (Fig.6.4 Table 6.2) indicated the presence of the psbA gene complementary region (4.5 Kb Xho I) juxtaposed to the inverted repeat unit in the large single copy region (Fig.6.6)

6.3.4 Gene for cytochrome f (cyt F)

Single stranded M13mp9 containing a 350 nucleotide Xho I - Bam HI fragment of the pea cytF gene, was employed to position its counterpart in soybean cpDNA. The double stranded probe constructed from the single strand copy, hybridized to 4.3 Kb Kpn I, 41.0 Kb Pvu II, 12.0 Kb Sac I and 5.6 Xho I, thereby establishing the presence of a region complementary to pea cytF sequences in the large single copy region approximately 3.0 Kb away from the atpA gene.

6.4 DISCUSSION

The localization of a gene is an important prerequisite to study its structural and functional organization. As a preface to the plastid gene expression studies, the chloroplast genes for various polypeptides have been positioned by hybridizing suitable gene probes from different plant species. The chloroplast DNA fragment

containing the identified gene could be used as a probe to faithfully localize the same gene on other chloroplast genomes because the plastid genes are highly conserved. The high degree of complementarity is evident from heterologous hybridization studies (Erion et al, 1981, Seyer et al, 1981, Palmer and Thompson, 1982) and the work presented here. This is the first report to position most of the plastid protein encoding genes on a single chloroplast DNA.

The various protein encoding genes which have been positioned include the gene for large subunit of ribulose biphosphate carboxylase oxygenase (*rbcL*), for alpha, beta, and epsilon subunits of CF component of ATPase (*atpA*, *atpB*, *atpE* respectively), for subunit III or the DCCD binding polypeptide of the CF component of ATPase (*atpH*), for 32 Kd thylakoid protein (*psbA*) and for cytochrome f (*cytF*).

In the soybean cpDNA, all the seven protein encoding genes are localized in the large single copy region.

The *psbA* gene in soybean, like other plant species is localized close to the end of one of the inverted repeat segments whereas the *rbc L* gene maps differently from other plants. In comparison to corn, cucumber, spinach and mustard, the carboxylase gene in soybean cpDNA is positioned nearer to the *psbA* gene (Fig. 2 in Link, 1981 and Fig. 4 in Palmer and Thompson, 1982). The soybean situation is similar to mung bean, a fellow member of the Leguminosae (Palmer et al, 1983)

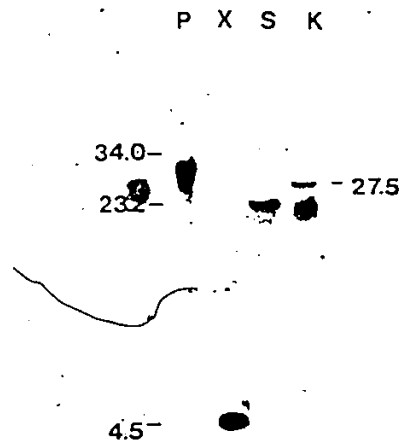


Figure 6.4 Hybridization of Nick-translated spinach psbA probe to nitrocellulose filter fixed fragments of soybean cpDNA with Kpn I (K), Pvu II (P), Sac I (S) and Xho I (X). The smearing in the Kpn I, Pvu II and Sac I lanes is due to the degradation of the hybridizing fragment.

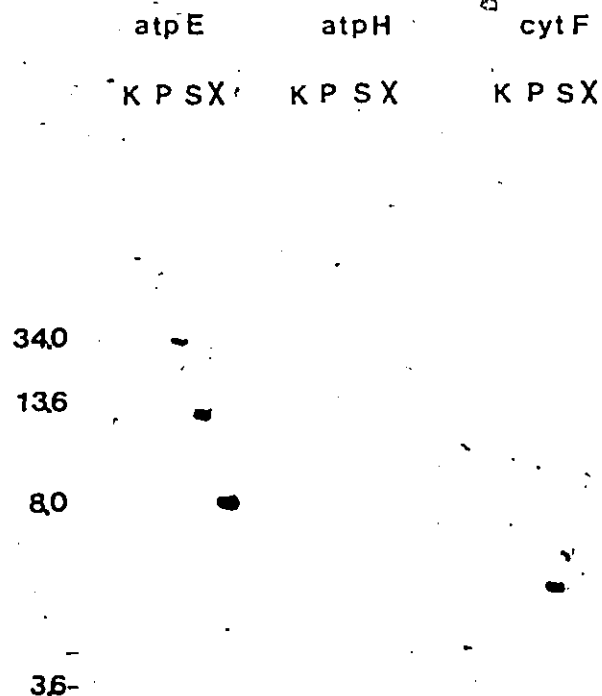


Figure 1. Mapping of the *atpE*, *atpH* and *cytF* genes on the soybean chloroplast DNA. ³²P labeled probes were hybridized against soybean DNA digested with Kpn I (K), Pvu II (P), Sac I (S), Xho I (X) fragments fractionated on agarose and transferred to nitro-cellulose filters.

The probes *atpE*, *atpH* are portions of wheat cpDNA cloned in single stranded M13 mp7 and *cytF* is part of pea cpDNA cloned in single stranded M13 mp9. The conversion of single stranded DNA into ³²P labeled double stranded molecule is described in methods.

Table 6.2 Positioning of the genes of ATP synthase and Cytochrome f by hybridization of ^{32}P labeled probes for the nitrocellulose filter bound soybean cpDNA.

Single digest	atpA	atpB	atpE	atpH	cytf
Kpn I	8.4	3.6	3.6	5.4	4.3
Pvu II	8.6	34.0	34.0	14.8	41.0
Sac I	5.6	23.2, 13.6	13.6	4.2	12.0
Xho I	27.0	8.0	8.0	27.0	5.6

Sizes of the fragments are in Kb.

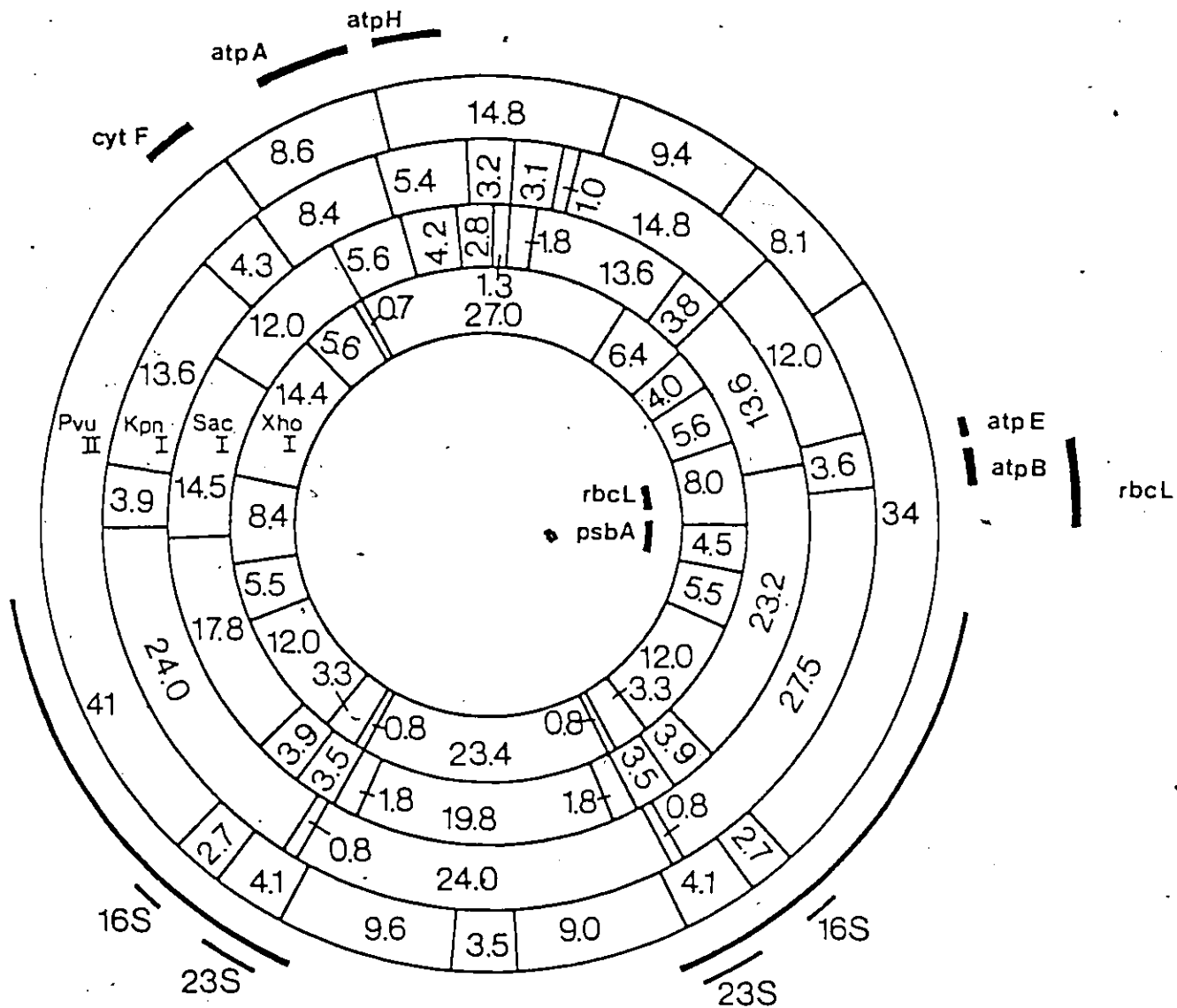


Figure 6.6 Localization of seven chloroplast protein genes on the physical map of the soybean cpDNA. The protein genes were positioned by Southern hybridization.
 atpA, atpB, atpE, atpH - genes for the subunits of ATP synthase.
 CytF - gene for cytochrome f
 psbA - gene for 32 Kd thylakoid membrane protein
 rbcL - gene for large subunit of RUBISCO

The hybridization pattern of *atpB* and *atpE* genes probes depict (Fig.6.6) that these genes are located in a region adjacent to *rbcL* gene, as has been noticed for pea, spinach and maize (Krebbers et al, 1982, Zurawski et al, 1982, 1983). In both

and spinach, the *atpB* and *atpE* genes are fused and are cotranscribed as a dicistronic message. The *atpB-atpE* gene block is separated from the *rbcL* by about 350 bp in corn and nearly 150 bp in spinach (Zurawski, 1982).

To find, if similar structural organization prevails in soybean, nucleotide analysis of the *rbcL* and *atpB-atpE* region is very essential.

The probes containing internal sequences of wheat *atpA* and *atpH* hybridized in the large single copy region, approximately 41 Kb and 36 Kb from the *atpB-atpE* gene block respectively. In comparison to soybean, the spinach *atpA* gene maps relatively proximal to the *psbA* encoding region (Westhoff et al, 1981). The radiolabeled probe carrying the front portion of the *cyt f* gene also hybridized to the large single copy region, thereby delineating the position of its complementary sequences on the soybean cpDNA.

The observed differences in the positioning of *rbcL*, *atpA*, *atpB* and *atpE* genes in soybean and spinach might be due to a big inversion in the large single copy region of either of the genomes with respect to each other as shown in the case of spinach and mung bean cpDNA (Palmer and

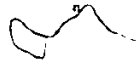
Thompson, 1982, Fig.4). Soybean and mung bean cpDNAs have been shown to be colinear in the gross sequence arrangement (Palmer et al, 1983). The exact effect of the structural rearrangements on the functional aspects of the genes or on the overall role of the plastid genome is yet to be detected. The comparative scrambling in the sequences of various chloroplast DNAs may be the result of evolution. It will be interesting to explore the possible relationship between the structural rearrangement of cpDNA and the functional organization of the organelle genome.

BIBLIOGRAPHY

- Bedbrook, J.R., Bogorad, L. 1976. Endonuclease recognition sites mapped on Zea mays chloroplast DNA. Proc. Natl. Acad. Sci. USA 73: 4309
- Bedbrook, J.R., Coen, D.M., Beaton, A.R., Bogorad, L., Rich, A., 1979. Localization of the single gene for the large subunit of ribulose biphosphate carboxylase on the maize chloroplast chromosome. J. Biol. Chem. 254: 905
- Bedbrook, J.R., Kolodner, R., Bogorad, L., 1977. Zea mays chloroplast ribosomal RNA genes are part of 22,000 base pair inverted repeat. Cell 11: 739
- Bedbrook, J.R., Link, G., Coen, D.M., Bogorad, L., Rich, A., 1978. Maize plastid gene expressed during photoregulated development. Proc. Natl. Acad. Sci. USA 75: 3060
- Birnboim, H.C., Doly, J., 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucl. Acids Res. 7: 1513
- Bohnert, H.J., Driesel, A., Herrmann, R.G., 1976. In genetics and biogenesis of chloroplast and mitochondria. (Ed - Bucher, Th., Neupert, W., Sebald, W., Werner, S.) North Holland Biomedical Press, Amsterdam. pp 629
- Bohnert, H.J., Driesel, A.J., Crouse, E.J., Gordon, K., Herrmann, R.G., Steinmetz, A., Mubumbila, M., Burkard, G., Weil, J.H., 1979. Presence of a transfer RNA gene in the spacer sequence between the 16S and 23S rRNA genes of spinach chloroplast DNA. FEBS Letters 103: 52
- Bohnert, H.J., Crouse, E.J., Schmitt, J.M., 1982. Organization and Expression of plastid genomes in Encyclopedia of Plant Physiol. : Nucleic Acids and proteins in plants. ed. Boulter, D., Parthier, B., 14B: 475 Springer, Heidelberg.
- Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heynecker, H.L., Boyer, H.W., Crosa, J.H., Falkow, S., 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene, 2: 95
- Bovenberg, W.A., Kool, A.J., Nijkamp, H.J.J., 1981. Isolation, Characterization and restriction endonuclease mapping of the Petunia hybrida chloroplast DNA. Nucl. Acids Res. 9: 503

- Bowman, C.M., Dyer, T.A., 1979. 4.5S ribonucleic acid, a ribosome component in the chloroplasts of flowering plants. *Biochem. J.* 183: 605
- Bowman, C.M., Koller, B., Delius, H., Dyer, T.A., 1981. A physical map of wheat chloroplast DNA showing the localization of the structural genes for the ribosomal RNAs and the large subunit of ribulose-1,5-bisphosphate carboxylase. *Mol. Gen. Genet.* 183: 93
- Boyer, H.W. and Roulland-Dussoix, R., 1969. A complementation analysis of the restriction and modification of DNA in Escherichia coli. *J. Mol. Biol.* 41: 459
- Britten, R.J., Kohne, D.E., 1968. Repeated sequences in DNA. *Science* 161: 529
- Britten, R.J., Graham, D.E., Neufeld, B.R., 1974. An analysis of repeating DNA sequences by reassociation, in *Methods in Enzymology*, Vol. 29E, Grossman, L., Moldave, K., Eds. Academic Press, N.Y., 1974. : 363
- Brosius, J., Palmer, M.L., Kennedy, P.J., Noller, H.F., 1978. Complete nucleotide sequence of a 16S ribosomal RNA gene from Escherichia coli. *Proc. Natl. Acad. Sci. USA*, 77: 201
- Brown, W.M., George, M. Jr., Wilson, A.C., 1979. Rapid Evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 76: 1967
- Brown, W.M. and Vinograd, J., 1974. Restriction endonuclease cleavage maps of animal mitochondrial DNAs. *Proc. Natl. Acad. Sci. USA* 71: 4617
- Burkard, G., Guillemaut, P., Weil, J.H., 1970. Comparative studies of the tRNAs and the aminoacyl-tRNA synthetases from the cytoplasm and the chloroplasts of Phaseolus vulgaris. *Biochem. Biophys. Acta* 224: 184
- Burkard, G., Steinmetz, A., Keller, M., Mubumbila, M., Crouse, E.J., Weil, J.H., 1982. Resolution of chloroplast tRNAs by two dimensional gel electrophoresis, in *Methods in chloroplast Molecular Biology* Eds. Edelman, M., Hallick, R.B., Chua, N.H. Elsevier/North Holland, Amsterdam.
- Cesareni, G., Muesing, M.A., Polisky, B., 1982. Control of Col El DNA replication: The rop gene product negatively affects transcription from the replication primer promoter. *Proc. Natl. Acad. Sci. USA* 79: 6313

- Chan, P.H. Wildman, S.G., 1972. Chloroplast DNA codes for the primary structure of the large subunit of Fraction I protein. *Biochim. Biophys. Acta* 277: 677
- Chu, N.M., Oishi, K.K., Tewari, K.K., 1981. Physical mapping of the pea chloroplast DNA and localization of the ribosomal RNA genes. *Plasmid* 6: 279
- Chua, N.H., Schmidt, G.W., 1978. Post translation transport into intact chloroplast of a precursor to a small subunit of ribulose-1,5-bisphosphate carboxylase. *Proc. Natl. Acad. Sci. USA* 76: 6110
- Chun, E.H.L., Vaughn, M.H., Rich, A., 1963. The isolation and characterization of DNA associated with chloroplast preparations. *J. Mol. Biol.* 7: 130
- Ciferri, O., 1978. The chloroplast DNA mystery. *TIBS* 3: 256.
- Coen, D.M., Bedbrook, J.R., Bogorad, L., Rich, A., 1977. Maize chloroplast DNA fragments encoding the large subunit of ribulose bisphosphate carboxylase. *Proc. Natl. Acad. Sci. USA* 74: 5487
- Collins, J., 1979. Escherichia coli plasmids packageable in vitro in bacteriophage particles. In *Methods of Enzymology*, Ed. Wu, R., Academic Press, New York, pp 555
- Collins, J., Brunning, H.J., 1978. Plasmids usable as gene-cloning vectors in an in vitro packaging by coliphage "cosmids". *Gene*. 4: 85
- Collins, J., Hohn, B., 1979. Cosmids: a type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage λ heads. *Proc. Natl. Acad. Sci. USA* 75: 4242
- Correns, C., Vererbungsversuche mit blass (gelb) grünen und bunt blattrigen sippen bei. Mirabilis urtica, und Lunaria. 1909a. *Zeit Induk. Abst. u. Vererbungs.* 1; 291
- Curtis, S.E., Haselkorn, R., 1983. Isolation and sequence of the gene for the large subunit of ribulose-1,5-bisphosphate carboxylase from the cyanobacterium Anabaena, 7120. *Proc. Natl. Acad. Sci. USA* 80: 1835
- Dahl, H.H., Flavell, R.A., Grosveld, F.G., 1981. The use of genomic libraries for isolation and study of eukaryotic genes. In *Genetic Engineering*, Vol 2., Ed. Williamson, R., Academic Press, London. pp 203
- Danna, K.J., Sack, G.H., Nathans, D., 1973. Studies on simian virus 40 DNA. VII. A cleavage map of the SV40 genome. *J. Mol. Biol.* 78: 363

- Darley-USmar, M.V., Fuller, S.D. 1981. Mr - values of mature subunits I and III of beef heart cytochrome c oxidase in relationship to nucleotide sequences of their genes. FEBS Letters 135: 164
- Deloris, H., Koller, B., 1980. Sequence homologies between Escherichia coli and chloroplast ribosomal DNA as seen by heteroduplex analysis. J. Mol., Biol., 142: 247
- Denhardt, D.T., 1966. A membrane-filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Comm. 23: 641
- Dobberstein, B., Blobel, G., Chua, N.H., 1977. In vitro synthesis and processing of a putative precursor for the small subunit of ribulose -1,5-bisphosphate carboxylase of Chlamydomonas reinhardtii. Proc. Natl. Acad. Sci. USA 74: 1082
- Dron, M., Rahire, M., Rochaix, J.D. 1982 Sequence of the chloroplast DNA region of Chlamydomonas reinhardtii containing the gene of the large subunit of the ribulose bisphosphate carboxylase and parts of its flanking genes. J. Mol. Biol. 162; 775
- Driesel, A.J., Crouse, E.J., Gordon, K., Bohnert, H.J., Herrmann, R.G., Stienmetz, A., Mubumbila, M., Keller, M., Burkard, G., Weil, J.H., 1979. Fractionation and identification of spinach chloroplast transfer RNAs and mapping of their genes on the restriction map of chloroplast DNA. Gene 6: 285
- Driesel, A.J., Speirs, J., Bohnert, H.J., 1980. Spinach chloroplast mRNA for a 32,000 dalton polypeptide. Size and localization on the physical map of the chloroplast DNA 1. Biochem. Biophys. Acta 610: 297
- Dugaiczky, A., Boyer, H.W., Goodman, H.M., 1975. Ligation of EcoR I endonuclease-generated DNA fragments into linear and circular structures. J. Mol. Biol. 96: 171
- Dyer, T.A., Bowman, C.M., Payne, P.I., 1976. The low-molecular weight RNAs of plant ribosomes: their structure, function and evolution. In Nucleic Acids and Protein Synthesis in Plants. ed. Bogorad, L., Weil, J.H. Plenum, New York : 135
- Dyer, T.A., Leech, R.M., 1968. Chloroplasts and cytoplasmic low-molecular weight ribonucleic acid components of the leaf of Vicia faba L. Biochem. J. 106: 689
- 

- Eaglesham, A.R.J., Ellis, R.J., 1974. Protein synthesis in chloroplasts II. Light-driven synthesis of membrane proteins by isolated pea chloroplasts. *Biochim. Biophys. Acta* 335: 396
- Edwards, K., Bedbrook, J., Dyer, T., Kossel, H., 1981. 4.5S rRNA from Zea mays chloroplast shares structural homology with the 3' end of prokaryotic 23S rRNA. *Biochem. Int.* 2:533
- Edwards, K., Kossel, H., 1981. The rRNA operon from Zea mays chloroplasts. Nucleotide sequence of 23S rDNA and its homology with *E. coli* 23S rDNA. *Nucl. Acids Res.* 9: 2853
- El-Gewely, M.R., Helling, R.B., Farmerie, W., Barnett, W.E., 1982. Localization of a phenylalanine tRNA gene on the physical map of the Euglena gracilis chloroplast genome. *Gene* 17: 337
- El-Gewely, R., Lomax, M.I., Lau, E.T., Helling, R.B., Farmerie, W., Barnett, W.E., 1981. A map of specific cleavage sites and tRNA genes in the chloroplast genome of Euglena gracilis. *Molec. Gen. Genet.* 181: 296
- Ellis, R.J., 1977. Protein synthesis by isolated chloroplasts. *Biochem. Biophys. Acta* 463: 185
- Ellis, R.J., 1978. Chloroplast proteins and their synthesis. In *Plant Proteins*, ed. Norton, G. London, Butterworths. 352 pp
- Ellis, R.J., 1981. Chloroplast proteins: Synthesis, Transport and Assembly. *Ann. Rev. Plant Physiol.* 32 : 111
- Erion, J.L., Tarnowski, J., Weissbach, H., Brot, N., 1981. Cloning, mapping, and in vitro transcription - translation of the gene for the large subunit of ribulose-1,5- bisphosphate carboxylase from spinach chloroplast. *Proc. Natl. Acad. Sci. USA* 78: 3459
- Feierabend, J., Schrader-Reichhardt, U., 1976. Biochemical differentiation of plastids and other organelles in rye leaves with high temperature-induced deficiency of plastid ribosomes. *Planta*. 120: 133.
- Feierabend, J., Wilder, G., 1978. Formation of the small in the absence of the large subunit of RUBP carboxylase in 70S ribosome-deficient rye leaves. *Arch. Biochem. Biophys.* 186: 283

- Feiss, M., Fischer, R.A., Crayton, M.A., Enger, C., 1977. Packaging of bacteriophage chromosome: Effect of chromosome length. *Virology*. 77: 281
- Fernald, M.L. 1950 In Gray's Manual of Botany. Eighth edition. American Book company, N.Y.
- Fluhr, R., Edelman, M., 1981a. Physical mapping of Nicotiana tabacum chloroplast DNA. *Molec.. Gen. Genet.* 181: 484
- Fluhr, R., Edelman, M., 1981b. Conservation of sequence arrangement among higher plant chloroplast DNAs: Molecular cross hybridization among the Solanaceae and between Nicotiana and Spinacia. *Nucl. Acids Res.* 9: 6841
- Gatenby, A.A., Castleton, J.A., Saul, M.W., 1981. Expression in *E. coli* of maize and chloroplast genes for large subunit of ribulose biphosphate carboxylase. *Nature* 291: 117
- Gatenby, A.A., Castleton, J.A., 1982. Amplification of Maize ribulose biphosphate carboxylase large subunit synthesis in *E. coli* by transcriptional fusion with the lambda N operon. *Molec. Gen. Genet.* 185: 424
- Gendel S., Straus, N.A., Pulleyblank, O., Williams, J.P., 1983. A novel shuttle cloning vector for the cyanobacteria. Anacystis nidulans. *FEMS Microbiol. Letters*. In press.
- Ginsburg, D., and Steitz, J.A., 1975. The 30S ribosomal precursor RNA From *Escherichia coli* a primary transcript containing 23S, 16S and 5S sequences. *J. Biol. Chem.* 250: 5647
- Godson, G.N., Vapnek, D., 1973. A simple method of preparing large amounts of x174 RFI supercoiled DNA. *Biochem. Biophys. Acta* 299: 519
- Gordon, K.H.J., Crouse, E.J., Bohnert, H.J., Herrmann, R.G., 1981a. Restriction endonuclease cleavage site map on chloroplast DNA from Oenothera parviflora. (Euoenothera Plastome IV). *Theor. Appl. Genet.* 59: 281
- Gordon, K.H.J., Crouse, E.J., Bohnert, H.J., Herrmann, R.G., 1981b. Physical mapping of differences in chloroplast DNA of the five wild-type plastomes in Oenothera subsection Euoenothera. *Theor. Appl. Genet.* 61: 373

- Graf, L., Roux, E., Stutz, E., Kossel, H., 1982. Nucleotide sequence of a *Euglena gracilis* chloroplast gene encoding for the 16S rRNA. Homologies to *E. coli* and *Zea mays* chloroplast 16S rRNA. *Nucleic Acid Research*. 10: 6369
- Graf, L., Kossel, H., Stutz, E., 1980. Sequencing of 16S - 23S spacer in a ribosomal RNA operon of *Euglena gracilis* chloroplast DNA reveals two tRNA genes. *Nature*. 286: 908
- Gray, P.W., Hallick, R.B., 1977. Restriction Endonuclease map of *Euglena gracilis* chloroplast DNA. *Biochemistry* 16: 1665
- Gray, P.W., Hallick, R.B., 1978. Physical mapping of the *Euglena gracilis* chloroplast DNA and ribosomal RNA gene region. *Biochemistry* 17: 284
- Gray, P. W., Hellick, R.B., 1979. Isolation of *Euglena gracilis* chloroplast 5S ribosomal RNA and mapping the 5S rRNA gene on chloroplast DNA. *Biochemistry*. 18: 1820
- Green, B.R., 1976. Covalently closed minicircular DNA associated with *Acetabularia* chloroplasts. *Biochem. Biophys. Acta* 447: 156
- Grossman, A., Bartlett, S., Chua, N.H. 1980. Energy-dependent uptake of cytoplasmically synthesized polypeptides by chloroplasts. *Nature* 285: 625
- Grunstein, M., Hogness, D., 1975. Colony hybridization: A method for isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Acad. Sci. USA* 72: 3961
- Haff, L.A., Bogorad, L., 1976. Hybridization of Maize chloroplast DNA with transfer ribonucleic acid. *Biochemistry* 15: 4105
- Hallick, R.B., 1983. Chloroplast DNA. In the biology of *Euglena*. ed Buetow, D.E., vol 4. Academic Press, N.Y.
- Hallick, R.B., Gray, P.W., Chelm, B.K., Rushlow, K.E., Orozio, E.M., 1979. *Euglena* Chloroplast DNA structure, gene mapping and RNA transcription. In chloroplast development. ed. Akoyounoglou, G.: 619. Elsevier, Amsterdam.
- Hartley, M.R., Ellis, R.J., 1973. Ribonucleic acid synthesis in Chloroplasts. *Biochem. J.* 134: 249
- Hartley, M.R. 1979. The synthesis and origin of chloroplast low molecular weight ribosomal ribonucleic acid in spinach. *Eur. J. Biochem.* 96: 311

- Hedberg, M.F., Huang, Y.S., Hommersand, M.H., 1981. Size of the chloroplast genome in Codium fragile Science 213: 445
- Helling, R.B., El-Gewely, M.R., Lomax, M.I., Baumgartner, J.E., Schwartzbach, S.D., Barnett, W.E., 1979. Organization of the chloroplast ribosomal RNA genes of Euglena gracilis bacillaris. Molec. Gen. Genet. 174: 1
- Herrmann, R.G., 1972. Do chromoplasts contain DNA? II. The isolation and characterization of DNA from chromoplasts, chloroplasts, mitochondria and nuclei of Narcissus. Protoplasma 74: 7
- Herrmann, R.G., Kowallik, K.V., 1970. Multiple amounts of DNA related to the size of chloroplasts. Protoplasma 69: 365
- Herrmann, R.G., Palta, H.K., Kowallik, K.V. 1980. Chloroplast DNA from three archegoniates. Planta 148: 319
- Holmes, D.S., Quigley, M., 1981. Rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. 114: 193
- Howe, C.J., Auffret, A.D., Doherty, A., Bowman, C.M., Dyer, T.A., Gray, J.C., 1982a. Localization and nucleotide sequence of the gene for the proton translocating subunit of wheat chloroplast ATP synthase. Proc. Natl. Acad. Sci. USA 79: 6903
- Howe, C.J., Bowman, C.M., Dyer, T.A., Gray, J.C., 1982b. Localization of wheat chloroplast genes for beta and epsilon subunits of ATP synthase. Molec. Gen. Genet. 186: 525
- Jenni, B., Stutz, E., 1978. Physical mapping of the ribosomal RNA region of Euglena gracilis chloroplast DNA. Eur. J. Biochem. 88: 127
- Kato, A., Shimado, H., Kusuda, M., Sugiura, M., 1981. The nucleotide sequence of two tRNA^{Asp} genes for tobacco chloroplasts. Nucl. Acids Res. 9: 5061
- Kashdan, M.A., Dudock, B.S., 1982. The gene for a spinach chloroplast Isoleucine has a Methionine anticodon. J. Biol. Chem. 257: 11191

- Keller, M., Burkard, H.J., Mubumbila, M. Gordon, K. Steinmetz, A., Heiser, D., Crouse, E.J., Weil, J.H., 1980. Transfer RNA genes associated with the 16S and 23S rRNA genes of Euglena chloroplast DNA. Biochem. Biophys. Res. Commun. 95 :47
- Ko, K., 1982 . MSc Thesis . Dept. of Botany . University of Toronto .
- Ko, K., Straus, N.A., Williams, J.P., 1983. Mapping the chloroplast DNA of Vicia faba. Curr. Genet. 7: 255
- Koch, W., Edwards, K., Kossel, H., 1981. Sequencing of the 16S - 23S spacer in a ribosomal RNA operon of Zea mays chloroplast DNA reveals two split tRNAs genes. Cell 25: 203
- Koller, B., Delius, H., 1980. Vicia faba chloroplast DNA has only one set of ribosomal RNA genes as shown by partial denaturation mapping and R-loop analyses. Molec. Gen. Genet. 178: 261
- Koller, B., Delius, H., 1982. A chloroplast DNA of Euglena gracilis with five complete rRNA operons and two extra 16S rRNA genes. Molec. Gen. Genet. 188: 305
- Kolodner, R., Tewari, K.K., 1975. The molecular size and conformation of chloroplast DNA from higher plants. Biochem. Biophys. Acta 402: 372
- Kolodner, R., Tewari, K.K., 1979. Inverted repeat in chloroplast DNA from higher plants. Proc. Natl. Acad. Sci. USA 76: 41
- Krebbers, E.T., Lorrinua, I.M., McIntosh, L., Bogorad, L., 1982. The maize chloroplast genes for p and e subunits of the photosynthetic coupling factor CF_1 are fused. Nucl. Acids Res. 10: 4985
- Kung, S.D., 1977. Expression of chloroplast genomes in higher plants. Ann. Rev. Plant Physiol. 28: 401
- Kuntz, M., Keller, M., Crouse, E.J., Burkard, G., Weil, J.H., 1982. Fractionation and identification of Euglena gracilis cytoplasm and chloroplastic tRNAs and mapping of tRNAs genes on chloroplast DNA. Current Genet. 6 : 63
- Kumar, A., Cocking, E.C., Bovenberg, W.A., Kool, A.J., 1982. Restriction endonuclease analysis of chloroplast DNA in interspecific somatic hybrids of Petunia. Theor. Appl. Genet. 62: 377

Link, G., 1981. Cloning and mapping of the chloroplast DNA sequences for two messenger RNAs from mustard (Sinapis alba L.) Nucl. Acid Res. 9: 3681

Link, G., Bogorad, L., 1980. Sizes, locations and directions of transcription of two genes on a cloned maize chloroplast DNA sequence. Proc. Natl. Acad. Sci. USA 77: 1832

Link, G., Chambers, S.E., Thompson, J.A., Falk, H., 1981. Size and physical organization of chloroplast DNA from mustard (Sinapis alba L.) Molec. Gen. Genet. 181: 454

Loening, V.E., Ingle, J., 1967. Diversity of RNA components in green plant tissues. Nature 215: 363

Lorimer, G.H., 1981. The carboxylation and oxygenation of ribulose-1,5- biphosphate. The primary events in photosynthesis and photorespiration. Ann. Rev. Plant Physiol. 1981. 32: 349

Lund, E., Dahlberg, J.E., Lindahl, L., Jaskunas, S.R., Dennis, P.P., Nomure, M., 1976. Transfer RNA genes between 16S and 23S rRNA genes in rRNA transcription units of *E. Coli*. Cell. 7: 165

Machatt, M.A., Ebel, J.P., Branlant, C., 1981. The 3' terminal region of bacterial 23S ribosomal RNA structure and homology with the 3' terminal region of eukaryotic 28S rRNA and with chloroplast 4.5, rRNA. Nucl. Acids Res. 9: 1533

Malnoe, P., Rochaix, J.D., 1978. Localization of 4S RNA genes on the chloroplast genome of Chlamydomonas reinhardtii. Molec. Gen. Genet. 166: 269

Malnoe, P., Rochaix, J.D., Chua, N.H., Spahr, P.F., 1979. Characterization of the gene and messenger RNA for the large subunit of ribulose 1,5. diphosphate carboxylose in Chlamydomonas reinhardtii. J. Mol. Biol. 133: 417

Mandel, M., Higa, A., 1970. Calcium dependent bacteriophage DNA infection. J. Mol. Biol. 53: 159

Maniatis, T., Fritsch, E.F., Sambrook, J., 1982. In molecular cloning: A laboratory manual. C.S.H. New York. pp 545

Maniatis T., Jeffrey, A., Klein, D.G., 1975. Nucleotide sequence of the rightward operator of phage. Proc. Natl. Acad. Sci. USA 72: 1184

- Mattoo, A.K., Pick, U., Hoffman-Falk, H., Edelman, M., 1981. The rapidly metabolized 32,000-dalton polypeptide of the chloroplast is the "proteinaceous shield" regulating photosystem II electron transport and mediating diuron herbicide sensitivity. *Proc. Natl. Acad. Sci. USA* 78: 1572
- Maizels, N., 1976 . Dictyostelium 17S , 25S and 5S rDNAs lie within a 38,000 base pair repeated unit. *Cell* 9: 431
- Mc Intosh, L., Poulsen, C., Bogorad, L.. 1980. Chloroplast gene sequence for the large subunit of ribulose biphosphate carboxylase in maize *Nature* 288: 556
- Mendolia-Morgenthaller, L.R., Morgenthaller, J.J., price, C.A., 1976. Synthesis of coupling factor CF₁ protein by isolated spinach chloroplasts. *FEBS letters* 62: 96
- Messing, J., Gronenborn, B., Muller-Hill, B., Hofschneider, P.W., 1977. *Proc. Natl. Acad. Sci. USA* 74: 3642
- Messing, J., 1980. In M13 cloning/sequencing. Instruction manual. Bethesda R.L. Inc.
- Messing, J., Crea, R., Seeburg, P.H., 1981. A system for shotgun DNA sequencing. *Nucl. Acids Res.* 9: 309
- Miller, K.R., Staehelin, L.A., 1976. Analysis of the thylakoid outer surface (coupling factor is limited to unstacked membrane regions). *J. Cell. Biol.* 68: 30
- Morgenthaller, J.J., Mendiola-Morgenthaller, L., 1976. Synthesis of soluble, thylakoid and envelope membrane proteins by spinach chloroplasts purified from gradients. *Arch. Biochem. Biophys.* 172: 51
- Mubumbila, M., Burkard, G., Keller, M., Steinmetz, A., Crouse, E., Weil, J.H., 1980. Hybridization of bean, spinach, maize and Euglena chloroplast transfer RNAs with homologous and heterologous chloroplast DNAs. *Biochem. Biophys. Acta* 609: 31
- Mubumbila, M., Gordon, K.H.J., Crouse, E.J., Burkard, G. and Weil, J.H..1983. Construction of the physical map of the chloroplast DNA of Phaseolus vulgaris and localization of ribosomal and transfer RNA genes. *Gene* 21 : 257
- Nelson, N., Nelson, H., Schatz, G., 1980. Biosynthesis and assembly of protein of the proton-translocating adenosine triphosphate complex from chloroplasts. *Proc. Natl. Acad. Sci. USA* 77: 1361

- Orozco, E.M. Jr and Hallick, R.B. 1982 Euglena gracilis chloroplast transfer RNA transcription units I Physical mapping of the transfer RNA gene loci. J.-Biol. Chem. 257 :3258
- Orozco, E.M. Jr., Rushlow, K.E., Dodd, J.R., Hallick, R.B., 1980. Euglena gracilis chloroplast ribosomal RNA transcription units. II. Nucleotide sequence homology between 16S - 23S ribosomal RNA spacer and the 16S ribosomal RNA leader regions. J. Biol. Chem. 255: 10997
- Osorio-Almeida, M.L., Guillemaut, P., Keith, G., Canaday, J., Weil, J.H., 1980. Primary structure of three leucine transfer RNAs from bean chloroplast. Biochim. Biophys. Res. Commun. 92: 102
- Pace, N.R., 1973. Structure and synthesis of the ribosomal ribonucleic acids of prokaryotes. Bacterial Rev. 37: 562
- Padmanabhan, U., Green, B.R., 1978. The Kinetic complexity of Acetabularia chloroplast DNA. Biochim. Biophys. Acta 521: 67
- Palmer, J.D., 1982. Physical and gene mapping of chloroplast DNA from Atriplex triangularis and Cucumis sativa. Nucl. Acids Res. 10: 1593
- Palmer, J.D., Singh, G.P., Pillay, D.T.N., 1983. Structure and sequence Evolution of three legume chloroplast DNAs. Molec. Gen. Genet. 190: 13
- Palmer, J.D., Stein, D.B., 1982. Chloroplast DNA from the fern Osmunda cinnamomea: Physical organization, gene localization and comparison to angiosperm chloroplast DNA. Current Genet. 5: 165
- Palmer, J.D., Thompson, W.F., 1981a. Rearrangements in the chloroplast genomes of mung bean and pea. Proc. Natl. Acad. Sci. USA 78: 5533
- Palmer, J.D., Thompson, W.F., 1981b. Clone banks of the mung bean, pea and spinach chloroplast genomes. Gene. 15: 21
- Palmer, J.D., Thompson, W.F., 1982. Chloroplast DNA rearrangements are more frequent when a large inverted repeat sequence is lost. Cell. 29: 537
- Palmer, J.D., Zamir, D., 1982. Chloroplast DNA evolution and phylogenetic relationships in Lycopersicon. Proc. Natl. Acad. Sci. USA 79: 5006

- Pillay, D.T.N., Singh, G.P., Guillemaut, P. 1983. (In preparation)
- Pillay, DTN and Cherry J.H. 1974. Changes in leucyl, seryl, and tyrosyl tRNAs in aging soybean cotyledons. Can. J. Bot. 52 :2499
- Pillay, D.T.N. and Gowda, S. 1981. Age related changes in transfer RNAs and synthetases in germinating soybean (Glycine max) cotyledons Gerontology. 27: 194
- Pfister, K., Steinback, K.E., Gardner, G., Arntzen, C.J., 1981. Photoaffinity labeling of an herbicide receptor protein in chloroplast membranes. Proc. Natl. Acad. Sci. USA 78: 981
- Poulsen, C., 1981. Comments on the structure and function of the large subunit of the enzyme ribulose biphosphate carboxylase-oxygenase. Carlsberg. Res. Commun. 46: 259
- Rawson, J.R.Y., Kushner, S.R., Vapnek, D., Alton, N.K., Boerma, C.L., 1978. Chloroplast ribosomal RNA genes in Euglena gracilis exist as three clustered tandem repeats. Gene. 3: 191
- Remaut, E., Stanssens, P., Fiers, W., 1981. Plasmid vectors for high efficiency expression controlled by the pL promoter of coliphage lambda. Gene 15: 81
- Rether, B., Bonnet, J. and Ebel, J.P., Studies on tRNA nucleotidyl transferase from Baker's yeast. IPurification of the enzyme Protection against thermal inactivation and inhibitor by several substrates. Eur. J. Biochem. 50 : 281
- Ris, H., Plaut, W., 1962. Ultrastructure of DNA containing areas in the chloroplast of Chlamydomonas. J. Cell. Biol. 13: 383
- Robert, R. 1980. Directory of restriction endonucleases. In Methods of Enzymology vol 65. ed Grossman, L. and Moldave, K. Academic Press, NY
- Rochaix, J.D., 1977. Restriction enzyme analysis of the chloroplast DNA of Chlamydomonas reinhardtii and construction of chloroplast DNA-plasmid hybrids in Acides Nucleiques et synthese de Proteines chez les vegetaux. ed. Bogorad, L., Weil, J.H., CNRS, Paris, 1977: 77
- Rochaix, J.D., 1978. Restriction endonuclease map of the chloroplast DNA of Chlamydomonas reinhardtii. J. Mol. Biol. 126: 597

- Rochaix, J.E., 1981. Organization, function and expression of the chloroplast 23S rRNA genes of Chlamydomonas reinhardtii: evolutionary and functional implications. J. Mol. Biol. 159: 383
- Rochaix, J.D., Darlix, J.L., 1982. Composite structure of the chloroplast 23S ribosomal RNA genes of Chlamydomonas reinhardtii. Evolutionary and Functional Implications. J. Mol. Biol. 159: 383
- Rochaix, J.D., Malnoe, P., 1978. Anatomy of the chloroplast ribosomal DNA of Chlamydomonas reinhardtii. Cell 15: 661
- Sato, S., Hutchison, C.A. III, Harris, J.I., 1977. A thermostable sequence-specific endonuclease from Thermus aquaticus. Proc. Natl. Acad. Sci. USA 74: 542
- Schwarz, Z., Jolly, S.O., Steinmetz, A., Bogorad, L. 1981. Overlapping divergent genes in the maize chloroplast chromosome in vitro transcription of the gene for tRNA. Proc. Natl. Acad. Sci. USA 78: 3423
- Schwarz, Z., Kossel, H., 1980. The primary structure of 16S rRNA from Zea mays chloroplast is homologous to E. coli 16S rRNA. Nature, 283: 739
- Sebald, W., Wachter, E., 1980. Amino acid sequence of the proteolipid subunit of ATP synthase from spinach chloroplasts. FEBS Letters 122: 307
- Seyer, P., Kowallik, K.V., Herrmann, R.G., 1981. A physical map of Nicotiana tabacum plastid DNA including the location of structural genes for ribosomal RNAs and the large subunit of ribulose biphosphate carboxylase/oxygenase. Curr. Genet. 3: 189
- Shen, G.F., Chen, K., Wu, M., Kung, S.D., 1982. Nicotiana chloroplast genome IV. N. accuminata has larger inverted repeats and genome size. Molec. Gen. Genet. 187: 61
- Shine, J., Seeburg, P.H., Martial, J.A., Baxter, J.D., Goodman, H.M., 1977. Construction and analysis of recombinant DNA for human chorionic somatomammotropin. Nature 270: 494
- Shinozaki, K., and Sugiura, M., 1982. The nucleotide sequence of the tobacco chloroplast gene for the large subunit of ribulose -1,5-bisphosphate carboxylase/oxygenase. Gene. 20: 91.

- Sinclair, D. and Pillay, D.T.N. 1980. Localization of tRNAs and Aminoacyl tRNA synthetases in cytoplasm, chloroplast and mitochondria of Glycine max, L. Z. Pflanzenphysiol Bd. 104 : 299
- Siegel, A., 1974. Gene amplification in plants. John Innis Symposium. 2nd. Norwich, England. Proc. 1974.
- Singh, G.P., Palmer, J.D., Pillay, D.T.N., 1982. A map of specific cleavage sites and rRNA genes in the chloroplast genome of soybean (Glycine max L). Plant Physiol. 69: 20a
- Sinsheimer, R.L., 1977. Recombinant DNA. Ann. Rev. Biochemistry, 46: 415 Annual Reviews, Palo Alto, Ca.
- Siu, C.H., Swift, H., Chiang, K.S., 1976. Characterization of cytoplasmic and nuclear genomes in the colorless alga Polytoma II. General Characterization of organelle nucleic acids. J. Cell. Biol. 69: 371
- Smith, G.E., Summers, M.D., 1980. The bidirectional transfer of DNA and RNA to nitrocellulose or diazobenzyloxymethyl-paper. Anal Biochem. 109: 123
- Smith, H.O., Brinster, M.L., 1976. A simple method for DNA restriction site mapping. Nucl. Acids Res. 3; 2387
- Southern, E., 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503
- Southern, E., 1979. Gel electrophoresis of restriction fragments in methods of Enzymology, 68. Ed. Wu, R., Academic Press, New York: 555.
- Spielmann, A., Ortiz, W., Stutz, E., 1983. The soybean chloroplast genome: Construction of a circular restriction sites map and location of DNA regions encoding the genes for rRNAs, the large subunit of the ribulose -1,5-bisphosphate carboxylase and the 32Kd protein of the photosystem II reaction center. Molec. Gen. Genet. 190: 1
- Steigler, G.L., Matthews, H.M., Bingham, S.E., Hallick, R.B., 1982. The gene for the large subunit of ribulose -1, 5-bisphosphate carboxylase in Euglena gracilis chloroplast DNA: location, polarity, cloning and evidence for an intervening sequence. Nucleic Acid Res. 10: 3427
- Steinmetz, A. 1979. These de Doctorat. Universite Louis Pasteur de Strasbourg, France.

- Steinmetz, A., Gubbins, E.J., Bogorad, L. 1982. The anticodon of the maize chloroplast gene for tRNA is split by a large intron. Nucl. Acids Res. 9: 5061
- Steinmetz, A. Krebbers, E.T., Schwarz, Z., Gubbins, E.J., Bogorad, L. 1983. Nucleotide sequence of five maize chloroplast transfer RNA genes and their flanking regions. J. Biol. Chem. 258: 5503
- Swamy, G.S. and Pillay, D.T.N., 1982. Characterization of Glycine Max cytoplasmic, chloroplastic and mitochondrial tRNAs and synthetases for phenylalanine tryptophase and turosine. Plant Sci. Letters 16: 157
- Sutcliffe, J.G., 1979. Complete nucleotide sequence of Escherichia coli Plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. 43: 77
- Takaiwa, F., Sugiura, M., 1980. Nucleotide sequence of the 4.5S and 5S ribosomal RNA genes from tobacco chloroplasts. Molec. Gen. Genet. 180: 1
- Takaiwa, F., Sugiura, M., 1982a. The complete nucleotide sequence of a 23S rRNA gene from tobacco chloroplasts. Eug. J. Biochem. 124: 13
- Takaiwa, F. and Sugiura, M., 1982b. Nucleotide sequence of the 16S - 23S spacer region in an rRNA gene cluster from tobacco chloroplast DNA. Nucl. Acids Res. 10: 2665
- Taniguchi, T., Fuji-Kuriyama, Y., Muramatsu, 1980. Molecular cloning of human interferon gene in Escherichia coli. Proc. Natl. Acad. Sci. USA 77: 5230
- Tohdoh, N., Sugiura, M., 1982. The complete nucleotide sequence of a 16S ribosomal RNA gene from tobacco chloroplasts. Gene 17: 213
- Tomika, N., Shirozaki, K., Sugiura, M., 1981. Molecular cloning and characterization of ribosomal RNA genes from a blue-green alga, Anacystis nidulans. Molec. Gen. Genet. 184: 359
- Thomas, P.S., 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA 77: 5201
- Thomas, J.R. and Tewari, K.K. 1974. Conservation of 70S ribosomal RNA genes in the chloroplast DNAs of higher plants. Proc. Natl. Acad. Sci. USA 71: 3147

- Thompson, J.A., Hansmann, P., Knoth, R., Link, G., Falk, H., 1981. Electron microscopical localization of the 23S and 16S rRNA genes within an inverted repeat for two chromoplast DNAs. *Curr. Genet.* 4: 25
- Twigg, A.J., and Sherratt, D., 1980. Trans-complementable copy-number mutants of plasmid ColE1. *Nature* 283: 216
- Ullrich, A., Shine, J., Chirgwin, R., Pictet, R., Tischer, E., Rutter, W.J., Goodman, H.M., 1977. Rat Insulin Genes: Construction of plasmids containing the coding sequences. *Science* 196: 1313
- Upholt, W.B., 1977. Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. *Nucleic Acid Research* 4: 1257
- Valenzuela, P., Venegas, A., Weinberg, F., Bishop, R., Rutter, W.J., 1978. Structure of yeast phenylalanine tRNA gene: an intervening DNA segment within the region coding for the tRNA. *Proc. Natl. Acad. Sci. USA* 75: 190
- van Ee, J.H., Vox, V.J., Planta, R.J., 1981. Physical map of chloroplast DNA of Spirodela oligorehiza: analysis by the restriction endonucleases Pst I, Xho I and Sac I. *Gene* 12: 191
- Wahl, G.M., Stern, M., Stark, G.R., 1979. Efficient transfer of large DNA fragments from agarose gels to diazobenzoxymethyl-paper and rapid hybridization by using dextran sulphate. *Proc. Natl. Acad. Sci. USA* 76: 3683
- Wallace, D.C., 1982, Structure and Evolution of Organelle Genomes. *Micro. Reviews.* 46: 208.
- Wensink, P.C., Finnegan, D.J., Donelson, J.E., Hogness, D.S. 1974. A system for mapping DNA sequences in the chromosome of Drosophila melanogaster. *Cell.* 10: 177
- Westhoff, P., Nelson, N., Biinemann, H., Herrmann, R.G., 1981. Localization of genes for coupling factor subunits on the spinach plastid chromosome. *Current Genetics.* 4: 109
- Wetmur, J.G., Davidson, N. Kinetics of renaturation of DNA. *J. Mol. Biol.* 31: 349
- Whitfield, P.R., Hermann, R.G., Bottomley, W., 1978. Mapping of the ribosomal RNA genes on spinach chloroplast DNA. *Nucleic Acid Research* 5: 1741

Whitfield, P.R., Leaver, C.J., Bottomley, W., Atchison, B.A., 1978. Low molecular weight (4.5S) ribonucleic acid in higher plant chloroplast ribosomes. *Biochem. J.* 175: 1103

Whitfield, P.R., Bottomley, W., 1980. Mapping of the gene for the large subunit of ribulose biphosphate carboxylase on spinach chloroplast DNA. *Biochem. Int.* 1: 172

Whitfield, P.R., Bottomley, W., 1983. Organization and structure of Chloroplast genes. *Ann. Rev. Plant. Physiol.* 34: 279

Willey, D.L., Huttly, A.K., Phillips, A.L., Gray, J.C. 1983. Localization of the gene for cytochrome f in pea chloroplast DNA. *Molec. Gen. Genet.* 189: 85

William, B.G., Blattner, F.R., 1979. Bacteriophage Lambda vectors for DNA cloning. In *Genetic Engineering: Principles and Methods*. Vol 2. (Ed. Setlow, J.K., Hollaender, A.) Plenum, N.Y. pp 289

Wurtz, E.A., Bueton, D.E., 1981. Intraspecific variation in the structural organization and redundancy of chloroplast ribosomal DNA cistrons in *Euglena gracilis*. *Curr. Genet.* 3: 181

Young, R.A., Macklis, R., Steitz, J.A., 1979. Sequence of the 16S-23S spacer region in two ribosomal operons of *E. coli*. *J. Biol. Chem.* 254: 3264

Zenke, G., Edwards, K., Langridge, P., Kossel, H., 1983. The rRNA operon from maize chloroplast: analysis of in vivo transcription products in relation to its structure. In *Cell Function and Differentiation*. ed. G. Akoyunoglu. In press.

Zhu, Y.S., Duvall, E.J., Lovett, P.S., Kung, S.D., 1982. Nicotiana chloroplast genome. V. Construction Mapping and expression of clone library of *N. otophora* chloroplast DNA. *Molec. Gen. Genet.* 187: 61

Zurawski, G., Bohnert, H.J., Whitfield, P.R., Bottomley, W., 1982. Nucleotide sequence of the gene for the 32,000-M_r thylakoid membrane protein from Spinacia oleraceae and Nicotiana debneyi predicts a totally conserved primary translation product of M_r 38,950. *Proc. Natl. Acad. Sci. USA* 79: 7699

Zurawski, G., Bottomley, W., Whitefield, P.R., 1982.

Structure of the genes for the β and ϵ subunits of spinach chloroplast ATPase indicates a dicistronic mRNA and an overlapping translation stop/start signal. Proc. Natl. Acad. Sci. USA 79: 6260

Zurawski, G., Bottomley, W., Whitfield, P.R., 1983.

Molecular analysis of adjacent transcriptional units from pea chloroplast DNA: *rbcL*, the gene for the large subunit of ribulose biphosphate carboxylase and *atpE*, the gene for the β and ϵ subunits of ATPase. In press.

Zurawski, G., Perot, B., Bottomley, W., Whitfield, P.R.,

1981. The structure of the gene for the large subunit of ribulose 1,5-bisphosphate carboxylase from spinach chloroplast DNA. Nucl. Acids Res. 9: 3251

VITA AUCTORIS

Date of Birth: September 24, 1957

Place of Birth: New Delhi, India

Education

1971-1974 Higher Secondary

Kendriya Vidyalaya
Delhi Cantonment

1974-1977 B.Sc. (Honours in Botany)

SGTB Khalsa College
University of Delhi, Delhi 110007

1977-1979 M.Sc. (Major-Genetics)

Dept. of Genetics
Punjab Agricultural University
Ludhiana - 141004

1979-1983 Ph.D.

Dept. of Biology
University of Windsor
Windsor, Ontario, Canada